The Dissertation Committee for Andrew Wallace Koch
certifies that this is the approved version of the following dissertation:

On Mutual Fund Herding

Committee:

Laura Starks, Supervisor

Andres Almazan

Aydogan Alti

Jonathan Cohn

Stefan Ruenzi

Clemens Sialm

On Mutual Fund Herding

by

Andrew Wallace Koch, B.S., M.B.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2011

On Mutual Fund Herding

Publication No. ________________

Andrew Wallace Koch, Ph.D.
The University of Texas at Austin, 2011

Supervisor: Laura Starks

This study examines several issues related to mutual fund herd behavior. First, a unifying and consistent framework for measuring herd behavior is developed. This framework generates portfolio-level measures for each fund manager over each quarter, and relates herd behavior to other aspects of portfolio dynamics. Simulations indicate significant and persistent non-random herd behavior. Second, mechanisms that potentially underly herd behavior are tested. Empirical results indicate that herding funds tend to i) change their holdings towards levels similar to peers, ii) have less experienced managers, and iii) underperform their peers. These results are consistent with a career concerns theory of herding. Third, the impact of mutual fund herding on stock liquidity is examined. Empirical results indicate that herd behavior can lead to correlation in stock-level liquidity.

Table of Contents

Chapter 1

Introduction

Herd behavior in a capital market setting is described by a group of investors moving money into or out of the same securities at the same time. Evidence over the past 40 years indicates that institutional investors may exhibit herd-like behavior.[1] More recently, research on investor herding has focused on a specific set of institutional investors, mutual fund managers. There is significant interest in understanding herd behavior among this group of professional investors for several reasons. First, the size of the mutual fund industry ($5.961 trillion in assets under management as of February 2011) suggests that common investment decisions among these managers may have considerable effects on asset prices and liquidity[2]. Second, the delegated nature of portfolio management in the mutual fund industry introduces several interesting theoretical explanations for their herd behavior. For example, do managers herd because it is in their own interest, or is it in the interest of shareholders? Third, mutual fund managers report their holdings quarterly. These publicly available data offer the econometrician an opportunity to measure and learn about herd behavior among fund managers both in the cross section and through time. The chapters that follow focus on these issues. Chapter 2 introduces a new framework to measure herding,

[1]See, for example, Kraus and Stoll (1972) for early evidence on herding among institutional investors.

[2]The size of the equity mutual fund industry as reported by Investment Company Institute, *Trends in Mutual Fund Investing*, March 30, 2011

Chapter 3 applies this framework in order to examine why fund managers herd, and Chapter 4 studies the impact that mutual fund herding may affect stock liquidity.

In Chapter 2 a framework for measuring investor herd behavior is proposed. Specifically, each portfolio is modeled as a physical entity moving through n-1 dimensional stock-space. The portfolio weights at the beginning of each quarter determine the fund's 'location', and as these weights change over time the fund moves through space. For each fund, the direction it is moving relative to its peers is measured over each quarter. This results in a fund, or portfolio-level, measure of herding over each quarter. In this way funds that trade in the same direction (herding funds) can be distinguished from funds whose trades resemble noise, and from those whose trades are contrarian.

The primary benefit of this framework is that additional metrics of portfolio evolution can be measured in a consistent manner, thus preserving any mechanical relationships. For example, it may be useful to examine how herd behavior relates to the similarity in portfolio holdings. Therefore the distance in stock-space between each fund and its peers is measured every quarter. This represents the similarity in the level of holdings. By using Euclidean geometry to measure not just the direction or distance, but any aspect of common portfolio evolution, all mechanical relationships between these metrics are preserved. In this way, the framework developed in this chapter provides a consistent and unifying method for measuring many aspects of portfolio dynamics.

This framework is used in Chapter 3 to test theories of investor herding. Fund managers may exhibit correlated trading because they are gathering information about stocks' future payoffs. If a group of managers has information for an overlapping set of securities, then they will tend to buy the same undervalued stocks, and sell the same overvalued stocks.

This is the informed trading explanation of herd behavior. On the other hand, managers might herd, not because they are creating value for shareholders by trading on information, but instead because it is personally costly for the manager to hold a portfolio that is too different from her peers. The idea that there may be 'safety in numbers' among mutual fund managers is the career concerns theory of herding. In other words, a manager who makes bad portfolio decisions is less likely to be fired if her peers also made bad decisions. However, if her decisions turn out to be wrong and she is alone in her actions, she will be deemed to be a poor quality manager and be fired. This can induce managers to focus less on creating value for shareholders, and more on herding in order to protect their personal careers.

In order to test these hypotheses, the direction a fund is trading relative to peers is computed in each quarter and for each fund. This direction is compared with the distance between funds in order to examine if funds are trading together in an effort to maintain similar portfolios as suggested by career concerns. Then an examination of the fund and manager characteristics that are associated with the direction measure is conducted.

Empirical evidence supports the career concerns hypothesis of herding. Fund managers that exhibit herd behavior tend to be trading towards the portfolio weights of their peers. These herding managers tend to be inexperienced relative to the managers whose trades are more independent. Finally, the subsequent performance of herding funds is low. These results are consistent with career concerns and difficult to reconcile with herding as the outcome of informed trading.

In Chapter 4, an examination of the impact of mutual fund herding on stock liquidity is conducted. Previous research documents a significant correlation in liquidity across stocks[3]. There are several possible reasons that stocks' liquidity might positively covary. There may be correlation across stocks in the adverse selection problem that generates a the liquidity premium. Or, there may be correlation in the costs of supplying liquidity from the perspective of the market maker (e.g. inventory costs). Chapter 5 tests an alternate hypothesis, that correlation in liquidity results from correlation in the demand for liquidity. That is, if there is a large group of investors that tend to own similar stocks, trade at similar times, and in similar directions, then these stocks are likely to experience liquidity shocks at similar points in time. Thus, stocks that are subject to correlated trading among mutual funds are likely to be those same stocks with correlated liquidity.

Evidence indicates that the liquidity of stocks which are heavily owned by mutual funds covary together. This effect is stronger when funds trade more, i.e. when ownership is weighted by fund turnover. The effect is also stronger during periods of high flows into or out of the mutual fund industry. These results support the hypothesis that correlated trading among mutual funds contributes to the observed common liquidity across stocks.

[3]See, for example, Chordia, Roll, and Subrahmanyam (2000), Hasbrouck and Seppi (2001), Huberman and Halka (2001), Eckbo and Norli (2002), Brockman, Chung and Perignon (2007), and Karolyi, Lee and vanDijk (2008) regarding commonality in liquidity

Chapter 2

A Framework for Measuring Herd Behavior

In this chapter I create a novel framework for measuring multiple aspects of herd behavior in a consistent and unifying way. The framework generates manager-level measures of common portfolio evolution. Each manager's portfolio is given a location in stock-space according to the portfolio weights. As the manager trades, the portfolio weights change and so does the location of the portfolio. For each manager and quarter I measure the direction the fund is moving in stock-space relative to its peers. This captures the similarity in both the sign and magnitude of the manager's trades with those of other funds. Secondly I measure the distance between a manager's portfolio from peers. This captures the similarity in the portfolio holdings at any given point in time. I describe the cross-sectional distribution of these measures and their persistence through time. Last, I compare managers' actual holdings and trades to a simulated system of portfolios evolving over time under the null hypothesis that trades are independent across managers.

2.1 Introduction and Literature Review

Investor herding is loosely defined as a group of investors trading in the same stocks at the same time in the same direction. Over the past four decades, researchers have offered several ways to measure investor herd behavior. The goal of this chapter is to introduce a new econometric approach to measuring investor herding. The framework developed within contributes to the literature in two important ways. First, I measure behavior at the fund-level, not stock-level as in previous literature. Second, the framework I employ allows for multiple aspects of portfolio evolution to be measured under the same methodology, thus preserving any mathematical relationships between metrics. This is useful if the researcher would like to examine any non-random relationships between herding and other portfolio dynamics. Third, other previously used portfolio metrics can be put into this framework, which pinpoints exactly how these seemingly unrelated measures used in previous literature are mathematically connected.

Kraus and Stoll (1972) is among the earliest empirical work on the topic. The authors use data on the trades of 229 institutional investors and construct the trade imbalance for each stock stock across all investors.[1] This is a simple way to capture the extent with which a group of investors are trading together. Several papers have used variants of the trade imbalance concept, such as changes in ownership, signed imbalances, or signed percentage imbalances.[2]

An important refinement is offered by Lakonishok et al. (1992). The authors define a measure of correlated trading that explicitly adjusts for the amount of correlated trading that would be expected under a null hypothesis of independent trades across investors. They calculate the fraction of funds trading in the same direction within a stock and quarter, and compare this across stocks (the LSV measure). Specifically, the LSV measure is defines as

[1]In a given stock-month the dollar net imbalance (DNI) is defined as $DNI = |TP - TS|$ where TP is the total dollars purchased by the 229 investors in that given stock-month, and TS is the total dollars sold.

[2]For example, see Wermers (1999), Sias (2004), Sias and Choi (2009), Pomorski (2009).

$LSV = |p_{i,t} - \bar{p}_t| - E|p_{i,t} - \bar{p}_t|$, where $p_{i,t}$ is the proportion of funds that are trading in stock i that are buyers of stock i in time t. The expectation term is calculated assuming $p_{i,t}$ is binomially distributed and the probability that a fund is a buyer is independent of other funds.

These measures of herding developed in previous literature have one general element in common; behavior is summarized at the stock-level. A stock-level measure of herding is the natural econometric approach if the researcher hopes to learn about effects of herding on prices, liquidity, or other variables that exist at the stock dimension. However, this approach to summarizing herding may not be best suited to address other related topics.

One such topic is the examination of the motivation(s) that may underly herding. Several papers have addressed this topic, and in doing so recognize that a fund, or manager-level measure of herding is best. The first of these papers, Grinblatt et al. (1995), generated a fund-level measure of herding by aggregating the LSV stock-level measure for each fund. The fund level measure, *FHM*, developed in Grinblatt et al (1995) is used to show that funds which tend to trade with the herd also trade on momentum.

Other papers have also used the *FHM* measure [e.g. Dass et al. (2008), Massa and Patgiri (2010)]. My paper contributes to these primarily by developing a framework for summarizing herd behavior at the fund-level in both trades and holdings, and showing how these behaviors are inter-related.[3]

Using mutual fund holdings from 1990-2006, I measure the direction a fund is changing

[3]I discuss in detail later how measuring behavior at the stock-level and then aggregating within each fund (such as *FHM*) can give quite different results from summarizing behavior at the fund-level (such as the measure developed later in this chapter).

its portfolio relative to its peers. This is the fund-level measure of herding. Additionally, I measure the similarity in the holdings of a fund and its peers which for reasons that will be clear later, is called *distance*. Numerous other metrics of portfolios can be summarized in the same framework, however these are the two upon which I focus. The measure of herding, *direction*, summarizes common portfolio dynamics for each fund, and as discussed it is most closely related to *FHM* developed by Grinblatt et al. (1995). The static measure of common portfolios, *distance*, is closely related to several previous measures of common portfolio holdings.

There are several papers that generate fund-level measures of the similarity in holdings. Two of these papers relate similarity to performance. Kacperczyk, Sialm and Zhang (2005) show that funds with industry weights similar to the market underperform those with dissimilar industry weights. Cremers and Petajisto (2009) measure the similarity in stock weights relative to an inferred benchmark and find that funds with weights similar to the benchmark underperform those with dissimilar stock weights. A third paper specifically relates similarity in holdings with career concerns. Chevalier and Ellison (1999) measure the similarity in industry holdings relative to an aggregate peer portfolio. They find that fund managers with holdings similar to peers tend to be inexperienced, younger managers who are fired more quickly conditional on performance. Taken together, the results in these papers are similar to the relationships I find between the similarity in holdings and performance or tenure. The primary contribution of my paper relative to these is that I connect the similarity in holdings with the literature on similar trading.

2.2 Methodology

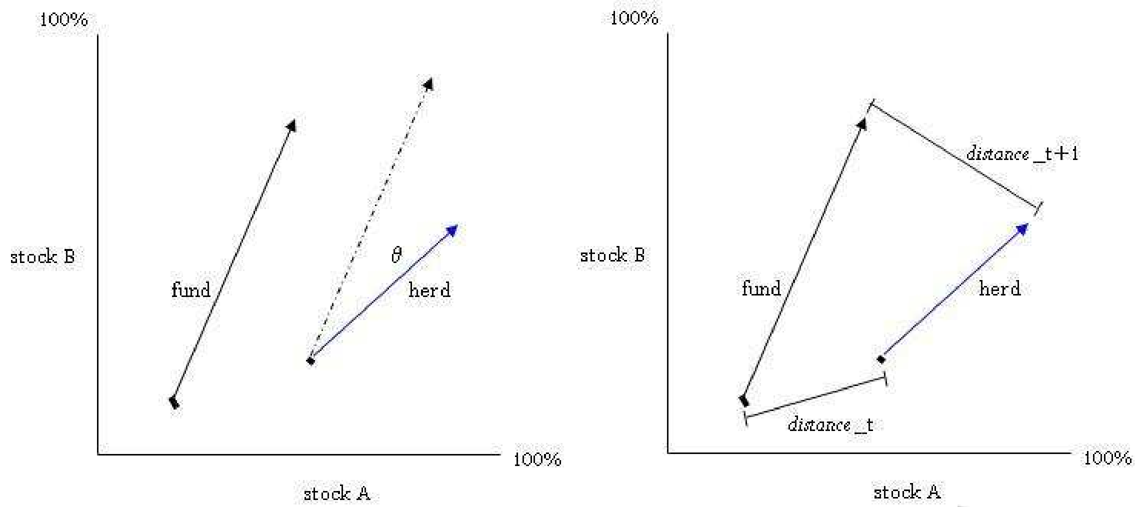
I summarize fund managers' investment behavior in two ways. Over each quarter I measure the tendency with which a given fund trades in the same direction as other peer funds. This is the fund-level analog to the traditional LSV stock-level measure of herding. The second fund-level measure of herd behavior is motivated by the career concerns mechanism. In addition to the tendency to trade together, I also measure the similarity in holdings between funds.

In order to measure these aspects of fund behavior in a consistent manner, I rely on a framework of physical herds such as those of birds or sheep. A herd of sheep might naturally be modeled by first indicating each entity's location on two-dimensional earth's surface. Then from each would be a vector pointing in the direction it is traveling, with length indicating its speed. One might examine the herding nature of a group by measuring the variance of the directions that the entities are moving. Or, one might examine the herding nature of a particular entity by asking if it is moving in the same direction as its peers. That is, if the angle between its vector and others' vectors is small, then it is herding.

Other characteristics in addition to the direction may be important to consider, and of course this will depend on context. For example, I will argue later that an entity's proximity to the herd may also be useful to distinguish the underlying mechanism(s) of herding. Is a fund located in the middle of the pack, or on the outskirts?

Each portfolio can be thought of as a physical entity whose location evolves over time. Portfolios are modeled as evolving through n-1 dimensions of stock space. A fund's location in stock space is determined by its portfolio weights. As the fund manager trades, the weights change and so the location of the portfolio evolves from one quarter to the next. The nature of the evolution of these portfolios through stock space can be summarized in a number of ways. I characterize two aspects that reflect herd behavior; the *direction* a fund is moving relative to its peers and the *distance* it is from its peers.

The direction a portfolio is moving relative to peers summarizes for each fund and quarter the degree with which the fund is trading with the herd. In the figure below, the direction the fund is moving is indicated by the angle θ between the fund and herd. In this 3-security example, both the fund and herd are increasing their portfolio weights in stock A and B, and by default decreasing their portfolio weighs in the remaining security (I discuss in the next section how I deal with this redundant dimension). In this particular example the fund is tilting its portfolio more strongly towards B than is the herd, and so they are not changing their portfolios in exactly the same direction. If the fund were buying and selling the same stocks that the herd was buying and selling, *and in the same proportion*, then θ would equal zero and the fund would be perfectly herding in portfolio weight changes. Conceptually, this is the measure of herd behavior closest to the stock level measure and is of primary interest.



The distance in stock space between a fund's portfolio and that of the herd reflects the similarity in portfolio weight levels. In the extreme, a fund that owns exactly the herd portfolio will be located in exactly the same place in stock space, and therefore have zero distance to the herd. As the fund's weights increasingly differ from the herd so will its distance in stock space. In the example above, the fund's portfolio is diverging, or becoming less similar to that of the herd. This highlights a case in which despite trading in a similar direction as the herd, the holdings become less similar over the quarter.

All aspects of portfolio evolution are measured with Euclidean geometry. This not only provides consistency across measures, but also relates closely to econometric counterparts. As I discuss later, the cosine of the angle between vectors is an un-centered, or non-Pearson correlation. The distance between a fund and the herd is the root sum of squared errors between their portfolio weights. These statistical measures are perhaps less intuitive in a physical sense but more commonly used in financial economics.

2.3 Results

I obtain quarterly mutual fund holdings data from Thompson for 1990 through 2006. I match holdings with stock and mutual fund data from CRSP using MFLinks. I retain all funds that are not identified as index funds per the index fund flag variable from CRSP and that have a benchmark as identified by Cremers and Petajisto (2009).[4] Stock holdings are adjusted for splits that occur between the record and file date. I exclude funds that do not report holdings in March, June, September, or December. This reduces the sample size as funds are not required to report holdings during these months, but has the benefit of ensuring common timing of information releases (to as fine a granularity as possible given the data). I require funds to report in consecutive quarters in order to estimate active portfolio weight changes. This reduces the sample size further, as funds are not required to report holdings quarterly throughout the entire time period. If a fund misses a report date it will periodically drop out of the sample.[5]

It is important to note that portfolio evolution through stock space is restricted. Portfolio weights must sum to 1 and so also their changes must sum to zero.[6] This means that one weight (or weight change) is a linear combination of the others. Therefore if we include all portfolio weights or changes in weights as independent observations we will overstate significance. To account for this, for each fund I drop the portion of its portfolio

[4]These data are provided by the authors at http://www.petajisto.net/data.html. These data are available for the 1990-2006 period. Later I also show results that do not rely on the availability of these data.

[5]There are 58,068 fund-quarter observations that merge to CRSP using MFLinks and have *activeshare*. Using only quarter-end report dates of March, June, September and December reduces the sample size to 46,328. Using only funds that report in consecutive quarters reduces the sample to 31,009.

[6]This is true for 92% of the fund-quarters in my sample - those whose value of holdings in Thompson match the asset values in CRSP.

in stocks with price less than five dollars.[7] Because results in this paper are largely cross-sectional, this adjustment should have little effect. But if we are to take the magnitudes of the measures of similarity seriously, then it would not be fair to treat each portfolio weight as a degree of freedom. So to be clear, I use all holdings to construct the weights and changes in weights, but before computing any similarities I drop the portion of a fund's portfolio in stocks with price less than 5.

All aspects of portfolio evolution are made in reference to an aggregate peer portfolio which I refer to as the herd or herd portfolio. I identify peer funds based on inferred benchmark using the data provided by Cremers and Petajisto. The authors infer a fund's benchmark based on the portion of its portfolio that is similar to each benchmark. The inferred benchmark is that to which the fund's portfolio is most similar. I require at least 10 funds to be in any given peer group and quarter in order to estimate similarities.

The method by which peers are identified is likely important for inference. In the next chapter, my goal is to distinguish evidence of agency problems from information, so a natural grouping would be an ex ante identifier of funds that are likely to observe the same signals or compete in the same job market. Using a holdings-based identifier such as the distance to a benchmark therefore seems appropriate. If the goal is to examine herd behavior relevant to asset prices, then it may make more sense to examine the mutual fund industry as a whole (because any correlated trading within a peer group may be washed out by any other peer group's correlated trading, thus little expected impact on prices). Or, if the goal

[7] Results are not dependent this choice. In earlier drafts I have used the portion of each fund's portfolio that cannot be matched to CRSP, which is nonzero for only 8% of the sample, as the redundant dimension. I choose the price=5 cutoff because the average total portfolio weight in these stocks is small and the gain in data manageability is large.

is to identify informed trades, one might try to use a finer grouping. This is exactly what is done in Pomorski (2009). He groups funds by family as these managers are quite likely to be trading on the same signals, and identifies informed trades as those that are common within the peer group.

The majority of results in this chapter and Chapter 3 use peer groupings by inferred benchmark. I include some robustness results using all funds industry-wide, and I have replicated core results using funds' objective code as reported by Thompson Reuters (unreported). Additional and likely informative groupings might be based on factor or cluster analysis. Importantly, the framework developed here is not restrictive in this sense.

At the beginning of each quarter, each fund's location in stock-space is defined by its portfolio weights. The vector of portfolio weights for fund f at quarter t is denoted $\mathbf{w}_{f,t}$, where each element i represents a security in the fund's portfolio, $w_{f,t,i} = \frac{shares_{f,t,i} * p_{t,i}}{\sum_i shares_{f,t,i} * p_{t,i}}$. $shares$ is the number of shares of stock i held by the fund as reported by Thompson and p is the stock price. For each fund $f = 1, 2, ..F$, the average peer fund portfolio has portfolio weights denoted by the vector $\mathbf{h}_{f,t} = \frac{1}{F-1} \sum_{-f} \mathbf{w}_{f,t}$ where $\sum_{-f}$ is the sum of all funds excluding fund f.

2.3.1 Common Portfolio Statics: *distance*

The Euclidean distance is determined by the Pythagorean Theorem, which is equivalent to root sum of squared errors between the two portfolios. This distance, which summarizes the similarity in portfolio levels is:

$$distance_{f,t} = ||\mathbf{w}_{f,t} - \mathbf{h}_{f,t}||$$

Related measures are *activeshare* [Cremers and Petajisto (2009)], *ICI* [Kacperczyk, Sialm and Zhang (2005)], and *SectorDeviation* [Chevalier and Ellison (1999)]. *distance* can be thought of as the shortest, or bee-line distance between the fund and herd, whereas *activeshare* is (one-half) the 'city block' distance to the benchmark. So *activeshare* differs from *distance* in two ways, the reference point and the mathematical calculation distance.[8]

The *ICI* measure in Kacperczyk, Sialm, and Zheng (2005) reflects portfolio industry concentration and is defined as the market-adjusted industry Herfindahl. A typical portfolio Herfindahl is the squared distance to the origin. This is equivalent to the squared Euclidean distance between the fund's portfolio and the market portfolio (in industry space). Similarly, the *SectorDeviation* measure in Chevalier and Ellison (1999) is also measured in industry space, but these authors use the Euclidean distance between the fund and the herd. In this way *SectorDeviation* is very similar to *distance*, the only difference being the dimensionality of space.

In summary, the *distance* measure is the root sum of squared errors between the portfolio weights of a given fund and the equal-weighted average weights of its peers. This is measured at the beginning and end of each quarter. In this way, the data is used to generate the similarity in the holdings of a fund relative to its peers in each cross section.

[8] The Euclidean distance that I use has the benefit of incorporating the nature of the differences, whereas activeshare (by design) does not. Activeshare treats two portfolio weight difference of 5% the same as ten portfolio weight differences of 1%. The Euclidean distance in the former case is $\sqrt{2 * 0.05^2} = 0.07$, but the distance in the latter case is $\sqrt{10 * 0.01^2} = 0.03$. Treating the latter case as less different is ideal because under this scenario the manager will have less tracking error. In this way the Euclidean distance incorporates how diversified the differences are.

2.3.2 Common Portfolio Dynamics: *direction*

In order to measure portfolio dynamics, I first adjust portfolio evolution for passive and active movements. Changes in prices are constantly reshaping stock-space, and so portfolios will move simply due to the passive realization of returns. While doing nothing may be a valid strategy, including both passive and active weight changes runs the risk of vastly overstating seemingly coordinated behavior. Therefore in each quarter I correct for the effects of returns on portfolio weights. The vector of weight changes from t, to $t+1$ is denoted $\Delta\mathbf{w}_{f,t}$ where each element i is defined

$$\Delta w_{f,t,i} = w_{f,t+1,i} - w_{f,t,i} * \frac{ret_{t,i}}{\sum_i w_{f,t,i} * ret_{i,t}},$$

and $ret_{i,t}$ is the return of stock i from t to $t+1$. For brevity I use $\mathbf{w}'_{f,t}$ to denote the vector of return correct portfolio weights (each element $w'_{f,t,i} = w_{f,t,i} * \frac{ret_{t,i}}{\sum_i w_{f,t,i} * ret_{i,t}}$).

So, each fund is modeled as a velocity vector $\Delta\mathbf{w}_{f,t}$ with location in stock space given by $\mathbf{w}_{f,t}$. This vector is corrected for the movement of the portfolio that would occur from passively realizing returns. Portfolio weights and changes are summarize in Table 2.1. I also report the number of funds in the sample, the number of stocks held by the herd, and the number of stocks held by a typical fund over a quarter.

To capture the similarity in portfolio weight changes, I measure the angle between the fund and herd. If the angle is zero, the portfolio is moving in exactly the same direction as the herd. The cosine function transforms the angle into a correlation, which is perhaps less intuitive but a more commonly used statistic. This correlation is non-Pearson, or uncentered, and is the metric I use to capture the similarity in portfolio changes:

$$direction_{f,t} = cos(\theta) = \frac{\Delta\mathbf{w}_{f,t} \bullet \Delta\mathbf{h}_{f,t}}{||\Delta\mathbf{w}_{f,t}||\ \ ||\Delta\mathbf{h}_{f,t}||}.$$

The angle between fund and herd vectors (or the cosine of the angle) is more appropriate than a typical correlation coefficient precisely because of the dropped redundant portfolio weight. Changes in weights have a reference state of zero and this would not be incorporated with a standard correlation coefficient. If all portfolio weights are used then by definition the intercept would be zero, but this would be incorrect because we would be treating each portfolio weight change as an independent observation when we know they must sum to 0. In other words, if we could use all weights then the intercept will be by definition zero, and a typical correlation coefficient would suffice. Because we need to drop a dimension we must also force the inclusion of the intercept into the correlation coefficient. This is exactly what is captured by the angle between the fund and herd vectors. The cosine of the angle between the two vectors is an uncentered correlation, which means that it incorporates the effect of a fund moving into or out of the dropped dimension without treating this movement as independent information. I summarize the computation of the various measures of portfolio evolution in Table 2.2.

I begin by describing the distribution of these measures of herd behavior. Panel A of Table 2.3 shows that the mean *direction* measure is 0.08 and median 0.051. The 10th and 90th percentiles are -0.03 and 0.22 respectively. On average, the distance between a fund and the herd is 0.13. The 10th and 90th percentiles are 0.08 and 0.18. For exposition I summarize these variables calculated using only non-zero beginning or ending quarter portfolio weights. Because a typical fund owns about 1/12th of the stocks owned by the herd (Table I), 11/12ths of a fund's weight changes will have zero correlation with peers.

This table shows that conditional on positive ownership sometime over the quarter, the average *direction* measure is about twice as high and the median is about 2.5 times as high. Portfolios are also not surprisingly closer to the herd when ignoring zero positions. The average *distance* in this case is 0.096. I show later that using this alternate definition does not affect results.

Panel B plots the distributions of *direction* and *distance*. The figure shows that *direction* has high positive skewness. There are a considerable number of funds with *direction* close to zero (none with exactly zero values). These are funds whose holdings overlap very little with those of their herd.[9] The distribution of *distance* is closer to normal but also has positive skewness. In the next section I examine the fund and manager characteristics that describe these distributions.

Panel C of Table 2.3 shows the persistence of *distance* and *direction*. In each quarter I rank funds into quintiles based on their measures of herding. I track these portfolios for up to 8 quarters and report the top-minus-bottom quintile spread. For example, the value of 0.067 in the bottom right cell indicates that funds put into quintiles based on *direction* 8 quarter in the past still exhibit a significant difference in their *direction* today. Therefore I find significant persistence in herd and contrarian behavior.

I use a variety of fund characteristics as controls in later tests. For each fund and quarter I collect from CRSP the fund's expense ratio, turnover ratio, date of fund inception, date that the current manager started at the fund, fund assets, and fund returns. I compute fund flows from CRSP returns and asset values, $flows_{f,t} = \frac{tna_{f,t+1} - ret_{f,t} * tna_{f,t}}{tna_{f,t}}$, and winsorize

[9] Core results are robust to the exclusion of funds with *direction* close to zero.

at 1%. *fund age* is the time in years between the current quarter and fund inception. Similarly *mgr tenure* is the time between the current quarter and the date the current manager started. I exclude observations in which *fund age* is less than zero. If *mgr tenure* is larger than *fund age* I set it equal to fund age. These adjustments are not critical to inference.[10] These variables are summarized in Table 2.4.

The average fund has $912 million in assets, expenses of 1.27%, and turnover ratio of 0.84. I sort funds in each quarter on *direction* and report average characteristics. Funds that strongly herd in trades (high *direction*) tend to be larger, although the relationship is not monotonic. These funds have high turnover, low flows and low subsequent raw returns. Funds that strongly herd in holdings (low *distance*) are also larger with low returns. There is a strong monotonic relationship between *direction* and *distance*. Funds that trade with the herd tend to have holdings very similar to the herd. This is the first evidence of a connection between herding in holdings and herding in trades.

2.3.3 Simulated Portfolio Evolution

Although not the focus of this research, an important issue in the investor herding literature is to compare herd behavior to that which we would expect due to chance. In order to address this, I simulate portfolio evolution using the empirical distribution of portfolio weight changes. For each herd (identified by benchmark) and for each quarter, I draw portfolio weight changes from the empirical distribution under the null hypothesis of inde-

[10] CRSP collects the date at which the current manager begins management directly from the source. In other words, if for example a fund is team managed, the manager tenure variable is determined by the date at which the fund itself reports the 'manager' took control of the fund. I am unable to distinguish if this represents the same team or team leader.

pendence. I do this for each quarter. In other words, at the beginning of each quarter, each fund begins with its actual portfolio weights and is assigned portfolio weight changes drawn from the sample of changes from that quarter and herd.

Several particularities of portfolios make this not quite straightforward. Each fund typically holds only a small fraction of the stocks owned by the herd and thus the majority of its weights are zero, and remain zero. Additionally, a number of positions are initiated over the quarter and others are closed out completely. There are likely several assumptions one could make in order to simulate realistic portfolio evolution. The specific design I describe below is relatively simple and matches statistics of actual portfolios relatively well. I have bootstrapped under several different assumptions and results are robust to the exact methodology.

For each benchmark group and in each quarter, I stratify the distribution of weight changes based on the size of the position at the beginning of the quarter. I draw portfolio weight changes with replacement from buckets based on the size of the initial portfolio weight. If the fund has zero weight in a stock at the beginning of the quarter, its weight change is drawn from the sample of actual weight changes across all funds and stocks in that herd and quarter in which the beginning period weight is zero. For non-zero initial positions, the weight change is drawn based on the level of ownership. I draw separately using 10, 25, 50, 75, and 90 percentile cutoffs. In these cases I draw the percentage of change in portfolio weight. This accounts for closing out positions (-100% change) without generating short positions. Last I normalize all funds' portfolios such that ending period weights sum to 1.

By drawing changes from the zero-initial-weight distribution and using percentage changes from non-zero distribution, I ensure that the number of stocks in the herd is unbiased.

In each quarter the number of positions closed and the number initiated match expectations.

Results are not sensitive to the structure of the bootstrapping. This design is in fact quite conservative, and results are generally stronger as assumptions are relaxed. By conditioning on initial weight levels I hardwire any herding related to the size of position.

This table summarizes various statistics for the actual and simulated evolution of portfolios. The output shows that the simulation matches the actual statistics quite well. I use these simulated portfolios to test for differences in herding between the portfolios that evolve independently (simulated) from the actual evolution of portfolios. In Table 2.3 Panel A I report the mean *direction* of 0.08 as significantly greater than the simulated sample at the 1% level.

2.3.4 Comparison with Other Herding Measures

To my knowledge, the only fund-level measure of herding that has been used in prior literature is the *FHM* measure developed in Grinblatt et al. (1995). In this section I compare *FHM* to *direction*.

The *FHM* measure is an aggregation within each fund of the *LSV* stock-level measure of herding. The first step to compute *FHM* is to sign the LSV measure, meaning buy herding is distinguished from sell herding. Second, for a given fund and quarter, the researcher adds across stocks all of the times that the fund trades with the herd, weighting each trade by the corresponding change in portfolio weight.[11] Specifically,

$$FHM = \sum_{i}^{N}(w_{i,3t} - w_{i,3t-3})(I_{i,t}LSV_{i,t} - E[I_{i,t}LSV_{i,t}]),$$

where w is the portfolio weight in stock i at time t, LSV is the stock-level herding measure defined in Lakonishok et al. (1992), and I is an indicator variable set to 0 if there is insignificant herd movement in the stock, 1 if there is significant herd buying and the fund is a buyer OR if there is significant herd selling and the fund is a seller, and -1 if the there is significant herd buying and the fund is a seller OR if there is significant herd selling and the fund is a buyer.

This measure reflects the extent with which a fund is changing its portfolio in the same direction as peers, when direction of trading is aggregated across all stocks. Conceptually this is very similar to *direction*, however *direction* adds important aspects to the *FHM* metric. First, *direction* allows for the distinction between active and passive portfolio decisions. That is, *FHM* may indicate a fund to be perfectly herding even the fund has made no trades. Through a simple adjustment to the portfolio weights *direction* makes this distinction. A second, more important difference is that *direction* accounts for the (similarity in) magnitudes of trades while *FHM* does not.

Consider the following example. Assume there are 100 mutual funds and three stocks, A, B and C. 99 mutual funds increase their portfolio weights by 20% each in A and B, and correspondingly decrease their portfolio weight in C by 40%. If the 100th mutual fund also changes its portfolio by +20%, +20%, and -40%, then its *direction* would equal 1, perfect

[11] For a detailed description of the aggregation of LSV herding measure within each fund and quarter, see Grinblatt, Titman and Wermers (1995).

herding. Any other portfolio change and the extent of herding would be less than perfect (*direction* < 1). However, the *FHM* measure would make no distinction between trades of +20%, +20%, and -40%, and trades of +0%, +40%, and -40%, for example. Also, a fund that moves all of its money into A, for example, would be considered to be a stronger herder than one that exactly mimicked the herd's portfolio weight changes. This is avoided by using variation in the data to directly measure behavior at the fund level, instead of first using variation to create a stock measure, then aggregating this at the fund.

2.4 Conclusion

In this chapter I develop a consistent and unifying framework for measuring herd behavior. The framework generates manager-level measures of common portfolio evolution. Multiple aspects of portfolio evolution are measured. These measures are compared to portfolio evolution simulated under the null hypothesis of independent trading.

Table 2.1
Summary of Fund Portfolios

I report the number of funds in the sample and the typical number of stocks they hold that can be matched to CRSP. I also report summary statistics on portfolio weights conditional on positive ownership at the beginning or end of the quarter. Importantly, I use all weights in all results going forward (included zero weights), however for illustrative purposes I summarize only non-zero weights in this table. *active* portfolio weight changes are changes in excess of that which would occur to passively holding the stock and realizing returns.

# funds	2089
# herds	19
avg # funds / qtr	461
avg # stocks held / qtr / fund	102
avg # stocks held / qtr / herd	1243

Portfolio characteristics
(conditional on positive ownership at beginning or end of quarter)

	Mean	Std Dev	Min	Max
weight	0.83%	1.05%	0.00%	100%
$\lvert\Delta$ weight$\rvert$	0.39%	0.61%	0.00%	100%
active $\lvert\Delta$ weight$\rvert$ end qtr	0.35%	0.61%	0.00%	100%

Table 2.2
Measuring Herding

The top portion of the table describes Measures of individual fund portfolios and how they evolve are described in the top panel. Below that are description and definitions of herding measures. Italics are scalars, boldface are vectors. Bounds for scalars are reported in the far right column. I also report a conceptual equivalence when possible.

Portfolio Measures	Description	Definition
$\mathbf{w}_{f,t}$	vector of portfolio weights at time t for fund f	$\langle w_{f,t,1}, w_{f,t,2}, ...\rangle$, where $w_{f,t,i} = \frac{shares_{f,t,i} * p_{t,i}}{tna_{f,t}}$
$\Delta\mathbf{w}_{f,t}$	vector of *active* weight changes where $w'_{f,t,i} = w_{f,t,i} * \frac{ret_{t,i}}{\sum_i w_{f,t,i} * ret_{i,t}}$	$\mathbf{w}_{f,t+1} - \mathbf{w}'_{f,t}$,
$\mathbf{h}_{f,t}$	herd portfolio weights, from perspective of fund f	$\frac{1}{F-1}\sum_{-f}\mathbf{w}_{f,t}$
$\Delta\mathbf{h}_{f,t}$	change in herd portfolio	$\frac{1}{F-1}\sum_{-f}\Delta\mathbf{w}_{f,t}$

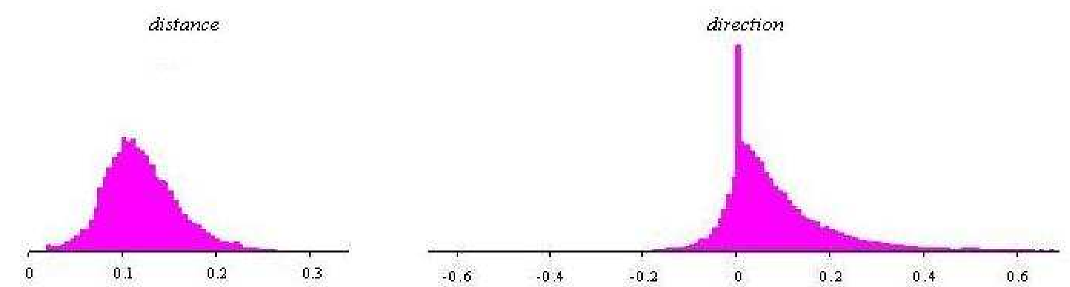
Herd Measures			Equivalence	bounds
$direction_{f,t}$	direction relative to herd	$\frac{\Delta\mathbf{w}_{f,t} \bullet \Delta\mathbf{h}_{f,t}}{\lvert\lvert\Delta\mathbf{w}_{f,t}\rvert\rvert\ \lvert\lvert\Delta\mathbf{h}_{f,t}\rvert\rvert}$	uncentered or non-Pearson correlation	[-1,1]
$distance_{f,t}$	distance between fund and herd	$\lvert\lvert\mathbf{w}_{f,t} - \mathbf{h}_{f,t}\rvert\rvert$	root sum of squared errors	$[0, \sqrt{2}]$

Table 2.3
Summary of Measures of Fund Behavior

Summary statistics on measures of portfolio evolution are reported in Panel A. *distance* reflects the similarity between a fund's holdings and those of the peer portfolio, $distance_{f,t} = ||\mathbf{w}_{f,t} - \mathbf{h}_{f,t}||$, which is the root sum squared errors between their portfolio weights. *direction* summarizes the similarity in portfolio weight changes, and equals the uncentered correlation between the fund and herd portfolio weight changes, $\frac{\Delta\mathbf{w}_{f,t}\bullet\Delta\mathbf{h}_{f,t}}{||\Delta\mathbf{w}_{f,t}||\ ||\Delta\mathbf{h}_{f,t}||}$. The stars in Panel A reflect significance of the *direction* variable compared to simulated portfolios (see Appendix). For exposition, at the bottom of Panel A I summarize the variables ignoring all portfolio weights that are zero. Panel B shows the empirical distributions of *distance* and *direction*. Panel C presents results on the persistence of *distance* and *direction*. In each quarter I sort funds into quintiles on either *distance* or *direction*. I track the average of these measures for each quintile for 8 quarters. In the table below I report the difference in the top and bottom quartile. For example, the bottom right cell shows that funds which were in the top quintile of *direction* 8 quarters ago now have *direction* that is 0.067 higher than funds that were in the bottom quartile 8 quarters ago. The t-statistic is computed using the time series of cross sectional average differences between the top and bottom quintiles.

Panel A: Summary Statistics	mean	median	std dev	10%	90%
distance	0.127	0.122	0.041	0.081	0.180
direction	0.080***	0.0512***	0.117	-0.026	0.224
distance [only non-zero weights]	0.096	0.089	0.032	0.063	0.137
direction [only non-zero weights]	0.157***	0.133***	0.204	-0.069	0.431

Panel B: Histograms



Panel C: Persistence - Top minus Bottom Quintile (5-1)

	quarter $t+$				
	+1	+2	+3	+4	+8
distance (5-1)	0.103	0.100	0.098	0.096	0.090
	(83.04)	(77.99)	(69.41)	(72.83)	(48.69)
direction (5-1)	0.010	0.084	0.087	0.082	0.067
	(12.61)	(12.56)	(12.10)	(13.36)	(12.39)

Table 2.4
Summary Statistics

The top panel summarizes these measures for the full sample, 1990-2006. The second panel reports mean fund characteristics by *direction*^{c} ranked quarterly. The third panel reports means by *convergence*^{c} quintile. Fund assets, expenses, turnover and returns are obtained from CRSP and merged using. Flows are computed from CRSP using returns and asset values. They are presented as a fraction of assets and are winsorized at 1 and 99%. activeshare is from Cremers and Petajisto (2009). return gap is from Kacperczyk, Sialm and Zheng (2008).

	Mean	Median	Std Dev
assets	912	189	3392
expenses	1.27%	0.0122	0.0045
turnover	0.84	0.62	0.82
lag flows	2.05%	-0.96%	17.25%
fund age	3.4	2.25	3.7
manager tenure	2.0	1.25	2.5
activeshare	0.78	0.81	0.17
return gap (bps)	-6	0	249
fundret (bps)	273	294	1079
lead fundret (bps)	254	294	1051

subsample means

			direction		
	Lo	1	2	3	Hi
assets	836	659	782	955	1352
expense ratio	1.28%	1.32%	1.30%	1.26%	1.21%
turnover	0.74	0.79	0.86	0.90	0.89
lag flows	2.10%	2.30%	2.28%	2.15%	1.47%
activeshare	0.81	0.84	0.80	0.75	0.68
return gap (bps)	-6	-2	-6	-3	-13
fundret (bps)	260	276	284	279	265
lead fundret (bps)	244	299	270	247	210
distance	0.141	0.137	0.127	0.121	0.112
direction	-0.037	0.018	0.058	0.116	0.246

			distance		
	Lo	1	2	3	Hi
assets	1409	1238	821	568	549
expense ratio	1.18%	1.24%	1.28%	1.32%	1.36%
turnover	0.83	0.92	0.89	0.86	0.69
lag flow	2.30%	1.88%	1.51%	2.21%	2.28%
activeshare	0.62	0.74	0.80	0.84	0.88
return gap (bps)	-2	-1	1	-8	-22
fundret (bps)	233	234	228	249	248
lead fundret (bps)	231	236	228	225	245
distance	0.078	0.105	0.122	0.143	0.190
direction	0.119	0.097	0.080	0.063	0.042

Table 2.5
Comparison of Actual and Simulated Portfolios

In each quarter I use actual starting portfolio weights and simulate portfolio evolution under the null hypothesis of independent weight changes. For each fund-quarter-stock I draw changes from the empirical distribution of changes from all peer funds in that quarter. The first two rows compare the change in weights and the ending portfolio weights for the non-zero portion of funds' portfolios, i.e. conditional on positive ownership at the beginning of the quarter. The next two rows summarize positions that are initiated, then I summarize positions that are closed out. The next last two columns summarize all portfolio weights, including all zero ownership positions. Last I show the measures of portfolio evolution in the simulated sample.

	actual		simulated	
	mean	std dev	mean	std dev
		$\|w_t > 0$		
Δw_t	-0.132%	0.586%	-0.131%	0.696%
w_{t+1}	0.851%	1.090%	0.852%	1.160%
		$\|w_t = 0$ and $w_{t+1} > 0$		
Δw_t	0.759%	0.797%	0.733%	0.772%
w_{t+1}	0.759%	0.797%	0.733%	0.772%
		$\|w_t > 0$ and $w_{t+1} = 0$		
Δw	-0.690%	0.830%	-0.750%	0.900%
w_{t+1}	0.000%	0.000%	0.000%	0.000%
		full sample		
Δw_t	-3.71e-4%	0.192%	-4.98e-4%	0.213%
w_{t+1}	0.064%	0.365%	0.064%	0.379%

	Herding measures: full sample			
direction	0.079	0.116	0.055	0.071

Chapter 3

An Empirical Examination of the Motives of Herd Behavior

In this chapter I apply the framework for measuring common portfolio evolution developed in Chapter 2 to data on quarterly mutual fund holdings over the 1990-2006 period in an effort to provide evidence on the motives that lead to herd behavior.

3.1 Introduction and Literature Review

Managers may herd because they are reacting to correlated information [e.g. Froot et al. (1992), Hirshleifer et al. (1994)] or they may be responding to the information in the trades of their peers [e.g. Banerjee (1992), Bikhchandani et al. (1992), Avery and Zemsky (1998)]. On the other hand, herd behavior may be the outcome of an agency problem between the manager and shareholders [e.g. Scharfstein and Stein (1990), Lakonishok et al. (1994), Chevalier and Ellison (1999)]. In this setting the manager's self interests, i.e. career concerns, may make it costly to deviate from peers. From the standpoint of the econometrician, the distinction between managers trading together because of similar information and managers trading together because of agency problems is difficult to determine.

I apply the herding measures developed in Chapter 2 to the data in order to distinguish the motives of herding. These measures are useful for a few reasons. First, by measuring herding at the fund-level I can examine fund qualities associated with herding, such as manager tenure and fund performance. Second, by examining similarity in holdings and similarity in trades in a single framework, I can test how these two behaviors relate to each other. This is useful because all theories of herding predict correlated trading, but career concerns is foremost a theory for correlated holdings.[1]

If fund managers herd because of career concerns, then herd behavior should be strongest among younger, inexperienced managers, and these funds should perform poorly in the cross section.[2] Also, if funds trade together because of career concerns then we would expect these funds to have portfolio levels that deviate very little from their peers. I find that funds with similar holdings have managers with short-tenure and low future performance. These same characteristics also describe funds with similar trades. Additionally, most correlated trading arises from the funds that hold very common portfolios and not from the funds that are willing to hold atypical portfolios. These results provide evidence that correlated trading is the outcome of the incentive for correlated holdings, and therefore that career concerns is an important driver of herd behavior.

To provide additional support, I test a conditional relationship between herding and performance. If the underperformance of funds with similar trades is driven by career concerns, then we would expect this effect to be strongest among funds that are also trading

[1] This concept is summarized by Scharfstein and Stein (1990) in the following quote regarding the bull market of 1987, "The consensus among professional money managers was that price levels were too high...However, few money managers were eager to sell. If the market did continue to go up, they were afraid of being perceived as lone fools. On the other hand, in the more likely event of a market decline, there would be comfort in numbers."

[2] Career concerns should result in lower performance in the cross section either because these managers push prices away from fundamentals [Grinblatt et al. (1995)], or because they are ignoring profitable information while those without career concerns are not [Scharfstein and Stein (1990)].

towards the holdings of their peers. I find that, unconditionally, funds that trade together underperform those with independent trades by 25 bps per quarter. Funds that are trading together and towards the holdings of their peers underperform by 76 bps per quarter.

3.2 Hypotheses Development

While correlated signals and career concerns are two of the more prominent theories of herding, there are other potential explanations. These include but are not limited to; changing preferences, irrationality, benchmark rebalancing, and liquidity-driven trading. I make an effort to control for some of these effects but otherwise do not speak directly to these alternative explanations. In this section I discuss in more detail how information or career concerns could lead to correlated trade behavior and how these different theories might uniquely manifest themselves in the data.

The focus of this paper is on contemporaneously correlated trading and holdings - where contemporaneous is defined as intra-quarter. Some theories of herding suggest that peers' trades need to be observable while others do not. First I discuss how the theories could result in contemporaneously correlated traded as measured using quarterly holdings data. Then I describe the fund and manager characteristics that we would expect to covary with herd behavior according to the theories. Last I describe predictions for how similarity in holdings should be related to similarity in trades.

3.2.1 Timing

In short, both information or career concerns could lead to either contemporaneous or lead-lag herding. It is clear that correlated information could lead to contemporaneously correlated trading. If fund managers obtain similar signals, then we would expect to find evidence of contemporaneous herding in trades as long as the managers are receiving their signals within the same quarter. Even if one set of investors consistently learn the information early in the quarter relative to other informed investors, herding is still likely [e.g. Hirshleifer et al. (1994)].

Informed trading can also show up as lead-lag herding. Generally, information cascades do not apply to settings in which prices adjust to information. However under some conditions, like those described in Avery and Zemsky (1998), a following group of funds may mimic the trades of a leading group because there is information in these trades. Again, lead-lag herding is not the focus of this chapter, but I do address it briefly.

Similarly, career concerns can generate either contemporaneous or cross-autocorrelated herding, although the mechanisms here are less clear. Interestingly, the literature that measures contemporaneously correlated trading often uses agency problems as potential motivation, but the connection between the two is somewhat incomplete.

There are several reasons to expect that career concerns may generate intra-quarter similarity in trades. First, a group of funds may have contemporaneously correlated trades - and thus be contemporaneously herding - because they are mimicking *something*. Fund manager might observe the holdings of their peers and trade towards them over the quarter. In a sense this is nothing more than rebalancing towards the herd portfolio. Because returns continuously reshape stock-space, managers need to trade in order to offset this effect of returns. If funds share deviations, which is what we would expect under career concerns, then the trades required to rebalance will be correlated. Importantly, the trades of these funds are correlated contemporaneously, and may or may not also be cross-autocorrelated with other funds' trades.

Or, it may be the case that trades do occur with a lag, we simply do not observe this due to data granularity. Some managers may trade together over the quarter because they observe each others' trades intra-quarter. Survey evidence compiled by Robert Shiller and John Pound in 1989 suggests this may be the case. The authors find that the majority of institutional investors claimed that their interest in new stock positions was driven by communication with other investment professionals. Importantly, these communications occurred contemporaneously with the acquisition of the position, "the time of maximum conversations for institutional investors tends to come after some purchases are made but while they are still purchasing" (p. 58).

Career concerns could also generate lead-lag herding. Career concerns predicts that fund managers want to mimic the holdings of their peers. This may require them to constantly mimic the trades of their peers, thus generating lead-lag herding.

3.2.2 Fund and Manager Characteristics

If funds herd because of their managers' career concerns we would expect herding to be strongest among managers with characteristics that positively covary with the agency problem between managers and shareholders. The agency problem that generates career concerns is likely the most severe among younger, less entrenched managers, and those who have recently performed poorly [Chevalier and Ellison(1999)]. Specifically, a manager early in his career has more at risk in terms of his human capital and so in the basic sense has higher career concerns. Similarly, the posterior probability that a seasoned manager is high ability will vary little from the ex ante expectation relative to that of a newly minted manager. On the other hand, if funds herd because their managers are informed, we would expect herding to be strongest among the more experienced fund managers (as these are the managers that have survived the job market) and these funds should have a strong performance track record.

I use manager tenure as a proxy for the magnitude of the agency problem between the manager and shareholders. This is obtained using the variable in CRSP that indicates the date the current manager took over management of the portfolio. As discussed earlier, manager tenure is defined as the difference in years between this date and the date at which the reported holdings are valid.[3]

3.2.3 Fund Performance

If managers ignore their information because of an incentive to maintain similar portfolios, these funds should underperform those willing to trade away from the herd. On the other hand if fund managers appear to herd because they are trading on the same signals, these funds should outperform those with managers that are not trading on the signal. I use the Carhart 4-factor model to adjust fund performance for risk and show that results are robust to using the 7 benchmark-based factors as defined in Cremers, Petajisto and Zitzewitz (2010).[4] Both of these include a momentum factor to account for the previously documented positive relation between herding and momentum trading [Grinblatt et al. (1995)].

[3] This date at which the current manager took over the portfolio is provided by CRSP. In cases of team managed funds, CRSP reports whatever date is reported to them by the mutual fund. I use this tenure variable in all cases and include an indicator variable for team managed funds.

[4] These data are provided by the authors at http://www.petajisto.net/data.html

3.2.4 Relation Between Holdings and Trades

In the results that follow, I provide evidence consistent with a setting in which funds with information tend to hold high *distance* portfolios, and funds hampered by agency problems tend to hold low *distance* portfolios. This is consistent with both the theory of career concerns [Scharfstein and Stein (1990), Zweibel (1995)] and previous empirical evidence [e.g. Chevalier and Ellison (1999) and Dass et al. (2009)]. Given this, then if funds trade together because of the incentive to hold similar portfolios, we would expect the correlated trading that we observe to arise from the funds that maintain portfolio holdings similar to the herd. On the other hand if correlated trading is the outcome of informed trading then we would expect these trades to arise predominantly among those funds willing to hold atypical portfolios. Additionally, the trades should serve the purpose of maintaining or increasing the similarity in holdings if funds trade together because of agency problems. There is no reason to expect that correlated trading due to information should result in holdings becoming more or less similar over time - sometimes trading on information may move their portfolios towards the holdings of their peers, and other times away.

More specifically, if career concerns drives herding then we would expect to find $\frac{\partial direction_t}{\partial distance_t} < 0$. And the trades of these funds should on average maintain or increase the similarity in holdings over time, $\frac{\partial distance_{t+n}}{\partial direction_t}|distance_t < 0$. If information drives herding then we should find $\frac{\partial direction_t}{\partial distance_t} > 0$. And the trades of these funds are not likely to have an effect of increasing or decreasing similarity in holdings over time, $\frac{\partial distance_{t+n}}{\partial direction_t}|distance_t = 0$

To summarize, either correlated signals or career concerns could lead to contemporaneously correlated trading.[5] If fund managers herd because of agency problems (information) then herding should be negatively (positively) related to tenure and negatively (positively) related to future performance. We would expect these herding funds to have similar(dissimilar) holdings and to be trading towards (independent of) the holdings of their peers.

3.3 Results

In this section I examine the lagged manager and fund characteristics that relate to herding in trades and herding in levels. I also examine the performance of these funds going forward. Then I examine the interaction between the two types of herd behaviors.

3.3.1 Similarity in Holdings: *distance*

The Euclidean distance between the fund and herd portfolio reflects the similarity in portfolio weight levels, and this is equivalent to the root sum of squared errors between the two vectors of portfolio weights at the beginning of the quarter. As discussed earlier, this is related to other static measures of active portfolio management, such as *activeshare* [Cremers and Petajisto (2009)], *ICI* [Kacperczyk et al. (2005)] and *SectorDeviation* [Chevalier and Ellison (1999)]. In this section I examine the lagged characteristics that describe the distribution of *distance* and how *distance* relates to subsequent fund performance.

The first column in Table 3.1 shows results from a pooled OLS specification with year and peer group effects. Funds with similar holdings are those that are close to the herd, i.e. low *distance*. Results indicate that these low *distance* funds tend to have low expenses

[5] Similarly, both could also lead to cross-autocorrelations. I examine this is a later section.

and higher turnover, and more assets under management. Importantly, after controlling for several fund characteristics, managers with short tenures are much more likely to have holdings similar to peers. The coefficient on manager tenure is 0.0022, indicating that a manager with 10 years less tenure holds a portfolio that is 0.022 closer to the herd. This is a considerable difference given an unconditional average distance of 0.127. Column 2 shows that all of these relationships are robust to median specification with bootstrapped standard errors.

The last two columns report odds ratios from a logit specification with dependent variable that equals one for high *distance* funds (column 3) or low *distance* funds (column 4). A manager with one additional year tenure is 1.19 times more likely to have a portfolio in the top quintile of *distance*, and 0.95 times likely to have a portfolio in the bottom quintile. In general, these results provide evidence that career concerns affect some managers' portfolio decisions. They are consistent with a setting in which fund managers with career concerns tend to hold portfolios similar to peers, and funds that trade on their information tend to end up holdings atypical portfolios.

Next I examine the relationship between the similarity in holdings and future performance. In Table 3.2 I report results from a Fama MacBeth specification regressing next quarter alpha on *distance* and controls. To estimate fund Carhart alpha I regress monthly fund returns from CRSP on zero cost market, size, book-to-market, and momentum portfolios. Alternatively, I also use a benchmark-based 7 factor model as defined in Cremers, Petajisto and Zitzewitz (2010), which I refer to as CPZ alpha.[6] I use 5 year rolling windows

[6]These data are provided by the authors at http://www.petajisto.net/data.html

and require at least 8 quarters (24 monthly observations) to estimate factor loadings. Using these factor loading I estimate the fund alpha over quarter t. I then regress the cross section of fund alphas in quarter t on *distance* and controls, both measured at quarter $t-1$. The additional explanatory variables are fund age, expenses, turnover, fund flows, and log of assets under management. All standard errors use Newey-West correction with maximum lags.

The dependent variable is next quarter's alpha in basis points. Column 2 reports the Fama MacBeth coefficients on lagged *distance* and controls. This shows that the relation between *distance* and performance is positive and significant both statistically and economically. A top-minus-bottom quintile portfolio on *distance* spreads gross returns $(0.19-0.078)*339 =$ 38 bps over the subsequent quarter, or 1.52% annually.

The next columns use dummies for high and low *distance* instead of the continuous variable. These results show that the positive relation is driven by the outperformance of funds with very different portfolios. Portfolios in the top quintile of *distance* outperform all other funds by 38 bps over the next quarter. In column 5 I include *activeshare* as an explanatory variable. The coefficient on *distance* drops moderately (from 339 in column 2 to 319 in column 5) and remains statistically and economically significant. The coefficient on *activeshare* is not distinguishable from zero. The last column shows that results are robust to an alternate risk adjustment that uses factors based on benchmark returns.

These results suggest that funds whose portfolio decisions are influenced by career concerns tend to be located in the middle of the pack, ie. low *distance*. And funds that seem to trade on valuable information have portfolios that are quite different from peers. Specifically, funds willing to hold atypical portfolios tend to be small funds with high expenses and their managers have long tenure. Their holdings seem to reflect skill or information as these funds outperform. In contrast, funds with portfolio weights that are closest to the average of their peers charge low fees and are managed by short-tenure managers. These funds underperform relative to their counterparts that hold atypical portfolios. Together, these results are consistent with Chevalier and Ellison (1999), and both Cremers and Petajisto (2009) and Kacperczyk et al. (2005). Chevalier and Ellison find that younger managers get fired more quickly and tend to hold industry weights similar to their peers. Cremers and Petajisto, and Kacperczyk et al. find a positive relationship between their measures of active management and performance.

3.3.2 Herding: *direction*

I use the cosine of the angle θ between the fund and herd vectors in order to measure the direction of trading. The cosine of θ is bounded [-1,1], where perfect herding is $direction = cos(0) = 1$, orthogonal trading is $direction = cos(90°) = 0$ and perfect contrarian trading is $direction = cos(180°) = -1$. In order to examine the characteristics that describe the distribution I regress *direction* on lagged characteristics and include time and peer group indicator variables. These results are presented in Table 3.3.

The specifications used here are the same as in Table 3.1. Funds with similar trades (high *direction*) have managers with short tenure. A manager with 10 years longer tenure has a measure of similarity in trades that is -0.03 lower. This is quite large compared to an unconditional mean of 0.08. Funds with similar trades also tend to have low expenses, high turnover, and high assets under management. These relationship hold using a median regression specification. The right two columns show that the tenure-*direction* relationship is driven by both the upper and lower end of the distribution. One additional year tenure means the manager is 0.95 times more likely to be in the top quintile of *direction* and 1.05 times more likely to be in the bottom quintile.

The relationship between performance and herding in trades is examined in Table 3.4. I use the same specifications as in Table 3.2. Fama MacBeth regressions show a significantly negative relationship between herding in trades and subsequent performance. After controlling for size, flows, turnover, expenses, and fund age, a top-bottom quintile portfolio spreads returns $(0.246 - -0.037) * -87 = -25$ bps per quarter, or 1.00% annually.

Interestingly, columns 3 and 4 show that the relationship is driven by both the underperformance of funds that trade together, and the outperformance of contrarian funds. Funds in the top quintile of similarity in trades underperform by 25 bps over the next quarter. Funds in the bottom quintile outperform by 11 bps. Both are statistically significant. These effects persist after controlling for *activeshare* and using CPZ benchmark alphas.

In summary, funds with similar trades tend to have managers with short tenure and they underperform relative to funds with more independent trades. These same relationships also describe funds that herd in holdings. The result that they share common characteristics suggests that they may be one in the same. I test this in the following section.

3.3.3 The Interaction between *direction* and *distance*

The results in the previous sections indicate that funds that herd in portfolio levels as well as those that herd in portfolio changes tend to be managed by individuals with short tenures, and these funds subsequently underperform other funds with more independent holdings and trades. These results are consistent with career concerns incentives to herd.

In this section I provide a few pieces of additional evidence supporting career concerns as a mechanism underlying both herding in holdings and herding in trades. Because career concerns is primarily a theory for correlated holdings and correlated trading is merely the outcome of the incentive to maintain similar portfolios, I test if these two herd behaviors are related. That is, do funds that trade together do so because of an incentive to maintain similar portfolios?

If correlated trading is due to career concerns we would expect high *direction* funds to have similar holdings (low *distance*), and to trade in such a way that this similarity is maintained. The first evidence of a connection between these two herd behaviors is provided in the summary statistics in Table 3.4. There is a very strong negative relationship between *distance* and *direction*, indicating that funds located near the herd are precisely those likely to exhibit correlated trading. The bottom quintile of *direction* has average *distance* equal to 0.141. This value among the top quintile of funds is 0.112 (Table 3.4). The t-statistic on the time series of the differences in cross-sectional sample means is 38.45.

A more detailed picture is presented in Figure 3.1 which shows the joint distribution. This shows that funds with similar trades (Hi *direction*) tend to have holdings that do not deviate much from their peers (Lo *distance*). Funds with independent/contrarian trades (Lo *direction*) tend to have holdings that deviate greatly from their peers (Hi *distance*).

These herding measures are endogenously determined by the same vector of beginning of quarter portfolio weights, and so it may be more interesting to examine not the location of funds that herd in trades, but where these funds are going. In Panel A, using a pooled OLS specification with quarter and benchmark effects and fund clusters, I regress *distance* at time $t + n$ on *direction* at time $t - 1$ and control for *distance* at time $t - 1$. Column 1 in

Table 3.5 shows that, conditional on beginning of quarter similarity in holdings, funds with similar trades tend to have portfolios very close to the herd two periods later (coefficient=-0.005, tstat=-3.54). These regressions include controls for fund return performance, flows, turnover, expenses and log assets. This effect persists for several quarters going forward.

To mitigate concern that this is mechanical, I show results repeating the analysis using simulated portfolios. For each fund I use the actual beginning of quarter portfolio weights and draw randomly with replacement portfolio weight changes from the contemporaneous empirical distribution of its peers. I draw from buckets based on the initial portfolio weight, e.g. zero ownership positions are drawn from other zero ownership positions (details in Appendix). The results using simulated portfolio evolution are reported in Panel B of Table 3.5. If anything, funds that trade together by chance tend to be trading away from the holdings of their peers[7]. These results show that funds that trade together tend to be trading towards the holdings of the herd.

Last, I examine a conditional *direction*-performance relationship. *direction* is a noisy reflection of herd behavior - because of both measurement error and chance. If the previously documented negative relationship between similarity in trades and performance is not spurious, but reflects agency problems associated with career concerns, then we would expect the underperformance to be driven by the funds that are also trading towards the holdings of the herd. In Table 3.6 I repeat the performance regressions from Table 3.4 and use subsamples based on whether or not the fund is trading towards the holding of its peers.

[7]My conjecture is that this is because portfolio weights are bounded [0,1]. If a fund has an initial weight of zero and its peers on average sell, then the fund cannot have a positively correlated trade in this stock, and its position wil generally become more similar to peers.

In each quarter I rank funds based on the change in the similarity in their holdings. Funds in the top quintile (strongly trading towards the holdings of their peers) in each quarter are used in columns 1 and 3, all others in columns 2 and 4. Specifically, I compute the percentage change in *distance* adjusted for the impact of returns on portfolio holdings.[8] Results show that the negative *direction*-performance relationship is driven by the funds that are trading towards the holdings of their peers. The impact of *direction* on performance is -271 among funds trading towards the holdings of their peers. This is quite large relative to the effect among the remainder of funds (-38), or relative to the effect among the full sample (-87, Column 2 of Table 3.4). Specifically the top-minus-bottom quintile spread is $(0.245 - -0.037) * -271 = 76$. This is three times the unconditional effect. This conditional relationship persists using the alternate benchmark-factor risk adjustment shown in columns 3 and 4.

3.4 Robustness Tests

In this section I show that the manager tenure and performance relationships are not sensitive to the specific definition of *direction*. In Table 3.7 I show that the negative relationship between tenure and herding persists across several fund-level measures of herding. The first three columns use herding measures computed within each benchmark group and columns 4-6 compute measures across all funds. In column 1 the dependent variable is *FHM* as defined in Grinblatt et al. (1995) and discussed in more detail in the previous chapter. It is calculated for each fund and quarter within each peer group. Column 2 uses

[8]The percent change in distance corrected for the impact of returns is $\frac{||\mathbf{w}_{f,t+1} - \mathbf{h}_{f,t+1}||}{||\mathbf{w}'_{f,t} - \mathbf{h}_{f,t}||}$.

FHM computed after defining all funds as peer funds. Column 3 uses $direction_{trd}$, which is the *direction* measure calculated after setting all portfolio weight changes to zero if the fund did not change the number of shares owned. Column 4 uses the standard *direction* measure computed defining all funds in the sample as peer funds.

The coefficient on mgr tenure is negative and significant in all cases. This indicates that the negative relationship between tenure and herding is robust to several measures of herding.

Table 3.8 repeats the performance regressions of Tables 3.4 and 3.6, but I use a number of different calculations of the two measures. The first variant is a moving average of *distance* and *direction* over the previous 4 quarters. Because these measures, particularly *direction* are noisy we may expect to better identify herding funds by those with consistently high *direction* or consistently low *distance*. Second, when controlling for the effect of returns I use returns without dividends[9]. Third, I use only the non-zero holdings of funds when computing the herding measures. This has the drawback of ignoring differences in the breadth of ownership but might arguably better reflecting active management of portfolios. Fourth, I measure behavior relative to a value-weighted herd vector. This would be a better approximation of the peer portfolio if we assume that the trades of larger funds are more informative or if fund managers are evaluated relative to large funds more often than relative to small funds. Last I compute the measures in industry space instead of stock space. This would be the appropriate methodology if information exists primarily at the industry level or if variation

[9]The effect that returns will passively have on portfolio weights depends on the assumption of what fund managers do with cash distributions. If managers reinvest into the stocks from which the dividends came then the correct return adjustment is to use returns with dividends. If fund managers pay out the dividends to shareholders then the correct return adjustment is to use returns without dividends.

in fund performance largely depends on industry, not stock, weights.

I report results on alternate definitions of *distance* in Panel A. The standard controls are included but not reported for brevity. *distance* remains a significant predictor of performance in all cases. The coefficient in column 5 is consistent with Kacperczyk et al. (2005) who find a positive relationship between a fund's market adjusted industry Herfindahl.

In Panel B I show results using alternate definitions of *direction*. The first column uses the average over the previous 4 quarters. Results indicate that this is likely to be a more precise measure of fund herding. The top-minus-bottom quintile spread in returns using *direction* is $(0.208 - -0.007) * -228 = -49$ bps per quarter, about twice the effect using the typical *direction* measure. The second column shows that results are robust to using returns without dividends to control for passive portfolio weight changes. The third column shows that restricting computations to use only non-zero holdings results in approximately the same estimated effect as the standard definition of *direction*. The top-bottom spread for *direction* is $(0.44 - -0.09) * -50 = -27$ bps using non-zero holdings. Last, I show results using measures computed in industry-space. I find no significant relationship between *direction* measured in industry space and performance.

3.5 Leaders and Followers

Results up to this point have focused on contemporaneous herding. Specifically, a fund's portfolio weight changes have been compared to its peers weight changes over the same quarter. In this section I report the descriptors and performance of leading and following funds. That is, what characteristics describe funds whose trades predict the next quarter trades of its peers? How do these leading funds subsequently perform. Similarly, which funds trades mimic those of its peers and how do these following funds perform?

These questions are important to understand because they add to our understanding of herding motives. The primary difference in the economic setting between contemporaneous vs. lead-lag behavior is observability. The career concerns motive of herding requires some type of observability. In the previous section regressions were designed to test if herding funds are trading towards the observable holdings of peers. In this section I test if there is a group of funds whose trades mimic the observable trades of peers. Importantly, these funds may or may not be contemporaneously herding.

The information motive of herding is slightly different in a lead-lag setting. If the herd is following a group of leaders for information reasons, then there needs to be a reason that this (observable) information is not priced. Generally these information cascade theories do not apply to capital markets settings in which prices adjust. However Avery and Zemsky (1998) describe an economic setting with dynamic prices in which information cascades can still occur. Furthermore, Pomorski (2009) shows empirically that mutual funds sometimes mimic the trades of previously outperforming funds, and that the trades of these following funds continue to perform well. So, despite the general non-applicability of information cascade theories to capital markets with dynamic prices, we have reason to believe that information cascades may describe mutual fund trading.

The identification of leading and following funds is straightforward. To identify leaders, for each fund I compute the *direction* measure using a fund's portfolio weight changes over quarter t and the weight changes of its peers over $t + 1$, specifically,

$$leader_{f,t} = \frac{\Delta \mathbf{w}_{f,t} \bullet \Delta \mathbf{h}_{f,t+1}}{||\Delta \mathbf{w}_{f,t}|| \ \ ||\Delta \mathbf{h}_{f,t+1}||}.$$

Similarly, to compute following funds I use a funds weight changes over t and those of its peers over $t-1$,

$$follower_{f,t} = \frac{\Delta \mathbf{w}_{f,t} \bullet \Delta \mathbf{h}_{f,t-1}}{||\Delta \mathbf{w}_{f,t}|| \ \ ||\Delta \mathbf{h}_{f,t-1}||}.$$

So, every fund has a leading and following measure, and to be clear, following funds need not be following leaders, and leaders need not be leading followers. Everything is measured relative to the herd, meaning leaders lead the herd, and followers follow the herd - so followers are quite indirectly following leaders as identified here. In the following section I report the fund and manager characteristics that describe the cross-sectional distributions of these metrics, and how these distributions relate to future performance.

3.5.1 Leaders and Followers: Results

First I examine the lagged fund and manager characteristics that describe the distributions of measures of leading and following. Table 3.9 reports coefficients from a pooled OLS specification, regressing *leader* or *follower* on lagged characteristics. All regressions include year and benchmark effects.

The first three columns show results using *leader* as the dependent variable. Column 1 shows that a fund is likely to lead others if it has performed well in the past, as the coefficient on perf perc (percentile rank of performance) is positive and significant. In columns 2 and 3, lagged measures of *leader* are included. These coefficients are insignificant, indicating that leading (or being followed) is not persistent after controlling for other manager and fund characteristics.

In columns 4-6 I examine the descriptors of *follower*. The coefficients in column 4 show that following funds tend to have performed poorly in the past and have low turnover. The coefficients on $follower_{t-1}$ and $follower_{t-4}$ in columns 5 and 6 respectively, are positive and significant. This indicates that there is a persistent group of following funds.

Evidence from these cross sectional regressions shows that leading funds have exhibiting strong prior performance, and following funds have delivered weak performance in the past. In Table 3.10 I examine the future performance of mutual funds as a function of leading and following behavior. The information cascade theory of lead-lag herding predicts that both leaders and followers are trading on profitable information. Career concerns predicts that following funds underperform in absolute terms (by pushing prices away from fundamentals) or relative to other funds (by ignoring information).

3.6 Conclusion

Correlated trading among mutual fund managers may be the outcome of informed trading or managerial career concerns. In order to distinguish these theories, I use a novel framework to generate two fund-level measures of herd behavior. I measure the direction a fund is trading relative to its peers. This measure is the portfolio-level analog to the stock-level measure of herding used in previous literature [Lakonishok et al. (1994)]. The second measure of herd behavior is the difference between the fund's holdings and its peers. This holdings-based measure is motivated by career concerns. While all theories of herding generally predict correlated trading, a career concerns theory of herding is primarily about correlated holdings.

I use these fund-level measures of herd behavior to distinguish career concerns motivations from informed trading in two ways. First I examine which fund qualities describe the cross sectional distributions of these measures. Second, I examine how the two herd behaviors relate to each other.

I find that funds with similar holdings have managers with less experience and these funds have low future performance. This is consistent with a career concerns theory of herding - managers have an incentive to ignore their signals and maintain similar holdings. I find similar relationship among funds that herd in trades. Correlated trading among funds is strongest among funds with short tenure managers, and these funds have low future performance. Consistent with this, I find that the majority of correlated trading arises from funds that hold portfolios that are very similar to their peers. I confirm that funds with similar trades are more likely to be trading towards the holdings of their peers.

Overall, results are consistent with a model of career concerns in which fund managers ignore information in order to stick together. Importantly, the findings in this paper suggest that at least some managers are informed. The negative herding-performance relationships suggests that fund managers who are willing to hold atypical portfolios or make uncommon trades do particularly well in the cross section. Herding, while presumably optimal for the manager does not seem to be particularly beneficial to the shareholder.

Table 3.1
Descriptors of *distance*

The dependent variable is *distance* which is a fund-quarter measure of the similarity in portfolio weight levels relative to an aggregate peer portfolio. Specifically,

$$distance_{f,t} = ||\mathbf{w}_{f,t} - \mathbf{h}_{f,t}||,$$

where $\mathbf{w}_{f,t}$ and $\mathbf{h}_{f,t}$ are vectors of beginning of quarter portfolio weights for the fund and herd, respectively. The first column uses a Pooled OLS specification with standard errors clustered at the fund. Column 2 reports results from a median regression with bootstrapped standard errors. Columns 3 and 4 report odds ratios from a conditional logit specification. The dependent variable in the third column equals one if *distance* is in the top quintile (ranked quarterly), zero otherwise. The dependent variable in column 4 equals one if *distance* is in the bottom quintile, zero otherwise. All regressions include year and benchmark indicator variables. All independent variables are lagged and obtained from CRSP. perf pctl is the relative performance of the fund over the previous quarter in percentile, where 1 is highest fund return, 0 is lowest. mgr tenure is the length of time in years in which manager has been with the fund and fund age is the time since inception, both in years.

	[1] Pooled OLS	[2] Median Regression	[3] Logit (odds ratios)	[4] Logit (odds ratios)
	distance	*distance*	hi *distance* dummy	lo *distance* dummy
mgr tenure$_{t-1}$	0.003***	0.002***	1.210***	0.908***
	(5.08)	(7.46)	(5.06)	(-2.58)
fund age$_{t-1}$	-0.001***	-0.001***	0.911***	0.996
	(-3.27)	(-4.98)	(-3.02)	(-0.15)
team mgmt$_{t-1}$	0.001	0.001	1.068	0.843
	(0.73)	(1.46)	(0.53)	(-1.40)
perf pctl$_{t-1}$	0.001	0.001	1.052	0.942
	(0.88)	(0.77)	(0.62)	(-0.86)
flows/assets$_{t-1}$	-0.002	0.002	1.358	1.102
	(-0.46)	(0.55)	(1.19)	(0.41)
turnover$_{t-1}$	-0.004**	-0.004***	0.686**	1.042
	(-2.50)	(-5.82)	(-2.47)	(0.62)
expenses$_{t-1}$	0.017***	0.017***	2.253***	0.385***
	(5.62)	(14.59)	(5.22)	(-4.90)
log(tna)$_{t-1}$	-5.1e-07***	-3.6e-07***	1.000*	1.000**
	(-2.64)	(-4.92)	(-1.92)	(2.26)
Observations	13405	13405	13400	13397
R-squared	0.10			
clustered s.e.'s	Y		Y	Y
bootstrapped s.e.'s		Y		

Table 3.2
Performance, *distance*

For each fund and quarter I estimate factor loadings using the previous 5 years of monthly returns from CRSP (at least 2 years required) and use these to compute the fund's 4-factor Carhart alpha and 7-factor CPZ [Cremers, Petajisto and Zitzewitz (2010)] benchmark-based alpha. I report results from a Fama-MacBeth specification regressing the fund's alpha (in bps) on lagged *distance* and controls. Standard errors use Newey-West correction with maximum lags.

	Carhart α_t (quarterly) (bps)					CPZ α_t
	[1]	[2]	[3]	[4]	[5]	[6]
$distance_{t-1}$	279.4***	339.1***			319.1***	275.2**
	(3.78)	(2.96)			(2.73)	(2.60)
hi *distance* dummy$_{t-1}$			37.57***			
			(3.29)			
lo *distance* dummy$_{t-1}$				2.86		
				(1.13)		
$activeshare_{t-1}$					11.94	
					(0.43)	
fund age$_{t-1}$		1.70*	1.75*	1.78*	1.47	-0.27
		(1.76)	(1.88)	(1.93)	(1.50)	(-0.27)
expenses$_{t-1}$		-11.33	-12.25*	-10.39	-12.27	0.79
		(-1.59)	(-1.87)	(-1.49)	(-1.55)	(0.17)
turnover$_{t-1}$		1.44	1.85	-1.28	0.65	7.22
		(0.24)	(0.33)	(-0.25)	(0.08)	(1.50)
flows/assets$_{t-1}$		50.47***	52.75	52.51**	49.93**	61.84***
		(2.75)	(0.79)	(2.47)	(2.61)	(4.59)
log(tna)$_{t-1}$		-5.74**	-6.05***	-8.09***	-5.27**	-3.52
		(-2.40)	(-2.83)	(-4.52)	(-2.05)	(-1.11)
Constant	-62.42***	-28.97	10.64	28.89**	-34.02	-41.34
	(-3.87)	(-1.58)	(0.79)	(2.05)	(-1.32)	(-1.59)
Observations	26604	16283	16283	16283	16283	16283
Number of groups	67	62	62	62	62	62

Table 3.3
Descriptors of *direction*

The dependent variable is *direction* which is a fund-quarter measure of the similarity in portfolio weight changes relative to an aggregate peer portfolio. Specifically,

$$direction_{f,t} = \frac{\Delta\mathbf{w}_{f,t} \bullet \Delta\mathbf{h}_{f,t}}{||\Delta\mathbf{w}_{f,t}|| \; ||\Delta\mathbf{h}_{f,t}||},$$

where $\Delta\mathbf{w}_{f,t}$ and $\Delta\mathbf{h}_{f,t}$ are vectors of return-adjusted portfolio weight changes for the fund and herd, respectively. All regressions include quarter and benchmark indicator variables. Column 2 uses a median regression specification. Columns 3 and 4 report odds ratios from a conditional logit specification. The dependent variable in the third column equals one if *direction* is in the top quintile (ranked quarterly and within each herd group), zero otherwise. The dependent variable in column 4 equals one if *direction* is in the bottom quintile, zero otherwise. All independent variables are lagged and obtained from CRSP. perf pctl is the relative performance of the fund over the previous quarter in percentile, where 1 is highest fund return, 0 is lowest. mgr tenure is the length of time in years in which manager has been with the fund and fundage is the time since inception, both in years.

	[1] Pooled OLS	[2] Median Regression	[3] Logit (odds ratios)	[4] Logit (odds ratios)
	direction	*direction*	hi *direction* dummy	lo *direction* dummy
mgr tenure$_{t-1}$	-0.0027***	-0.0022***	0.942***	1.046***
	(-3.24)	(-3.70)	(-3.18)	(2.83)
fund age$_{t-1}$	4.0e-05	0.0007	1.000	1.001
	(0.06)	(1.40)	(0.02)	(0.09)
team mgmt$_{t-1}$	-0.004	-0.002	0.894	0.983
	(-1.10)	(-0.97)	(-1.50)	(-0.29)
perf pctl$_{t-1}$	0.002	-0.003	1.010	0.965
	(0.57)	(-1.09)	(0.14)	(-0.47)
flows/assets$_{t-1}$	-0.007	-0.009	0.671**	1.111
	(-0.76)	(-1.51)	(-2.06)	(0.58)
turnover$_{t-1}$	0.010***	0.012***	1.190***	0.800***
	(4.07)	(7.75)	(3.89)	(-3.33)
expenses$_{t-1}$	-0.028***	-0.019***	0.569***	1.246***
	(-5.61)	(-7.94)	(-5.11)	(2.70)
log(tna)$_{t-1}$	1.5e-06***	1.8e-06***	1.000***	1.000
	(2.95)	(5.43)	(2.62)	(-1.41)
Observations	13405	13405	13400	13395
R-squared	0.16			
clustered s.e.'s	Y		Y	Y
bootstrapped s.e.'s		Y		

Table 3.4
Performance, *direction*

For each fund and quarter I estimate factor loadings using the previous 5 years of monthly returns from CRSP (at least 2 years required) and use these to compute the fund's 4-factor Carhart alpha and 7-factor CPZ [Cremers, Petajisto and Zitzewitz (2010)] alpha. I report results from a Fama-MacBeth specification regressing the fund's alpha (in bps) on lagged *direction* and controls. Control variables are fund age, expenses, turnover, fund flows, and log assets. Standard errors use Newey-West correction with maximum lags.

	Carhart α_t (quarterly) (bps)					CPZ α_t
	[1]	[2]	[3]	[4]	[5]	[6]
$direction_{t-1}$	-79.56***	-86.84**			-78.07**	-86.11**
	(-2.71)	(-2.48)			(-2.32)	(-3.64)
hi *direction* dummy$_{t-1}$			-25.01**			
			(-2.48)			
lo *direction* dummy$_{t-1}$				10.60**		
				(2.44)		
$activeshare_{t-1}$					47.00*	
					(1.77)	
fund age$_{t-1}$		1.54*	1.69*	1.51*	1.35	-0.52
		(1.68)	(1.86)	(1.66)	(1.40)	(-0.74)
expenses$_{t-1}$		-11.10	-11.27	-9.11	-13.92*	-0.56
		(-1.45)	(-1.59)	(-1.24)	(-1.70)	(-0.05)
turnover$_{t-1}$		-0.15	-0.25	-0.30	-1.41	5.87
		(-0.10)	(-0.12)	(-0.13)	(-0.35)	(1.45)
flows/assets$_{t-1}$		59.05***	60.19***	54.12**	57.91***	71.75***
		(3.14)	(3.19)	(2.58)	(2.88)	(5.20)
log(tna)$_{t-1}$		-7.30***	-7.21***	-7.71***	-6.74***	-4.22*
		(-4.37)	(-4.21)	(-4.60)	(-3.38)	(-1.80)
Constant	-19.33	32.24*	29.78**	24.24*	-3.57	10.69
	(-1.31)	(2.18)	(2.07)	(1.77)	(-0.23)	(0.85)
Observations	26604	16283	16283	16283	16283	16283
Number of groups	67	62	62	62	62	62

Figure 3.1
Joint Distribution of *distance* and *direction*

The figure below is a histogram of the joint distribution of *distance* and *direction*. The z-axis reports the frequency of observations across 5x5 independent sorts (each quarter) on the two herding measures. Hi(Lo) *distance* represents funds with very dissimilar(similar) holdings. Hi(Lo) *direction* represents funds with very similar(dissimilar) trades.

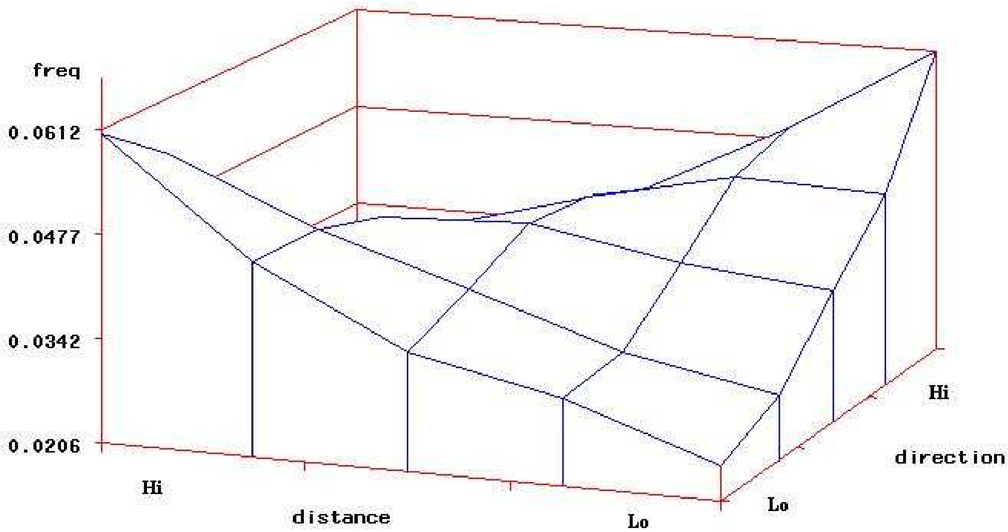


Table 3.5
Interaction of *distance* and *direction*

I regress *distance* at time $t + n$ on direction over the $t - 1$ quarter. All regressions include time and benchmark indicator variables, fund clusters, and controls. Control variables are fund performance percentile, flows as a percentage of assets, turnover, expenses and log assets, all measured at t-1. Panel A *distance* and *direction* values calculated from the data. For comparison, Panel B shows results using values calculated from the bootstrap simulation. I simulate portfolio evolution under the null of (conditionally) independent trading. Specifically in each quarter I use actual beginning period portfolio weights and draw weight changes from the empirical distribution [details in the Appendix]. I use these simulated portfolio weight changes to calculate the resulting *direction* over the $t - 1$ quarter and the end of quarter *distance*.

	$distance_{t+1}$	$distance_{t+2}$	$distance_{t+3}$	$distance_{t+4}$
Panel A: Actual				
$direction_{t-1}$	-0.00452***	-0.00533***	-0.00517***	-0.00502**
	(-3.54)	(-3.14)	(-2.78)	(-2.42)
$distance_{t-1}$	0.918***	0.885***	0.862***	0.844***
	(100.12)	(69.21)	(62.23)	(59.79)
Constant	0.0110**	0.0109*	0.0191***	0.0146***
	(2.33)	(1.72)	(3.47)	(2.96)
Controls	√	√	√	√
Observations	13645	11380	10679	9923
R-squared	0.88	0.84	0.80	0.77
Panel B: Simulated Portfolios				
$direction_{t-1}$	0.0101*	0.00498	0.00742	0.0192**
	(1.74)	(0.71)	(1.19)	(2.25)
$distance_{t-1}$	0.735***	0.708***	0.694***	0.682***
	(72.00)	(55.07)	(49.66)	(46.39)
Constant	0.0733***	0.0411***	0.0477***	0.0495***
	(8.32)	(7.60)	(7.91)	(9.77)
Controls	√	√	√	√
Observations	13645	11380	10679	9923
R-squared	0.66	0.62	0.60	0.56

Table 3.6
Performance, Interaction of *distance* and *direction*

I repeat Fama MacBeth regressions of performance on *direction* and controls on subsamples. Columns 1 and 3 include funds that are strongly trading towards the holdings of their peers and columns 2 and 4 use all others. Specifically I rank funds at the end of each quarter on the change in *distance* (corrected for returns).

	Carhart α_t (quarterly) (bps)		CPZ α_t (quarterly) (bps)	
	[1] trading towards peers' holdings	[2] all others	[3] trading towards peers' holdings	[4] all others
$direction_{t-1}$	-271**	-38*	-318***	-50*
	(-2.03)	(-1.86)	(-2.71)	(-1.95)
fund age$_{t-1}$	-1.2	2*	-1.7	-0.1
	(-0.77)	(1.89)	(-1.34)	(-0.13)
expenses$_{t-1}$	5.9	-6	23.7	-48
	(0.35)	(-0.98)	(1.38)	(-0.09)
flows/assets$_{t-1}$	65***	58***	65*	81***
	(2.93)	(3.87)	(1.77)	(5.93)
log(tna)$_{t-1}$	-9.8**	-7***	-7.8	-2.7
	(-2.63)	(-5.03)	(-1.64)	(-1.36)
turnover$_{t-1}$	6.6	-5	11	3.6
	(0.78)	(-0.92)	(0.74)	(1.07)
Constant	38	21	24	-3.5
	(1.33)	(1.24)	(0.59)	(-0.39)
Observations	3257	13026	3257	13026
Number of groups	62	62	62	62

Table 3.7
Robustness, Characteristics of Herding Funds

This table reports coefficients from regressing one of a variety of measures of fund herding on lagged manager and fund characteristics. The first three columns use herding measures computed within each benchmark group and columns 4-6 compute measures across all funds. In column 1 the dependent variable is *FHM* as defined in Grinblatt et al. (1995) calculated for each fund and quarter within each peer group. Column 2 uses *FHM* computed after defining all funds as peer funds. Column 3 uses $direction_{trd}$ which is the *direction* measure calculated after setting all portfolio weight changes to zero if the fund did not change the number of shares owned. Column 4 uses the standard *direction* measure but computed after defining all funds as peer funds. All regressions include quarter and benchmark indicator variables. All independent variables are lagged and obtained from CRSP. perf pctl is the relative performance of the fund over the previous quarter in percentile, where 1 is highest fund return, 0 is lowest. mgr tenure is the length of time in years in which manager has been with the fund and fund age is the time since inception, both in years. All regressions include year and benchmark indicator variables and standard errors are clustered at the fund.

	[1]	[2]	[3]	[4]
	FHM	$FHM_{allfunds}$	$direction_{trd}$	$direction_{allfunds}$
mgr tenure$_{t-1}$	-0.0004***	-0.0002***	-0.0019***	-0.0016***
	(-2.82)	(-3.09)	(-2.99)	(-2.62)
fund age$_{t-1}$	4.3e-05	0.0001**	0.0006	0.0007
	(0.43)	(2.03)	(1.10)	(1.26)
team$_{t-1}$	-0.0005	-0.0003	-0.0062**	-0.0054*
	(-0.95)	(-1.02)	(-2.12)	(-1.88)
perf pctl$_{t-1}$	0.0006	-0.0008**	0.0038	0.0030
	(0.87)	(-2.18)	(1.38)	(1.15)
flows/assets$_{t-1}$	-0.0021	-0.0019**	-0.0146**	-0.0090
	(-1.31)	(-2.00)	(-2.08)	(-1.35)
turnover$_{t-1}$	0.0095***	0.0042***	0.0064***	0.0055***
	(10.16)	(8.13)	(3.05)	(2.82)
expenses$_{t-1}$	-4.2e-05	5.3e-05	-0.0147***	-0.0151***
	(-0.05)	(0.12)	(-3.70)	(-3.91)
log(tna)$_{t-1}$	-6.5e-08	-2.0e-09	1.2e-06***	1.2e-06***
	(-1.32)	(-0.09)	(2.81)	(2.63)
Observations	13405	13405	13399	13405
R-squared	0.25	0.31	0.17	0.18

Table 3.8
Robustness, Performance of Herding Funds

For each fund and quarter I estimate factor loadings using the previous 5 years of monthly returns from CRSP (at least 2 years required) and use these to compute the fund's 4-factor Carhart alpha and 7-factor CPZ [Cremers, Petajisto and Zitzewitz (2010)] benchmark-based alpha. I results from Fama-MacBeth specification regressing the fund's alpha (in bps) on lagged measures of herding and log assets. I calculate *direction* and *distance* using 5 alternate definitions. Results on alternate *distance* definitions are in Panel A, and *direction* in Panel B. The first row using the 4 quarter moving avg of the measures. The second row uses ex-dividend returns to adjust portfolio weight changes for the passive effect of returns. The third row uses measures after restricting stock-space to those stocks held by the fund at either the beginning or end of the quarter. The fourth row value-weights peer funds to construct the herd vector. The fifth row reports results using dimensions defined by 48 industries as determined by Fama and French (1997). Columns 1-5 use Carhart risk adjustment and columns 6-10 use Cremers et al. risk adjustment. All specifications include control variables log assets, fund age, expenses, turnover, and fund flows. Standard errors use Newey-West correction with maximum lags.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
	Carhart α_t (quarterly) (bps)					CPZ α_t (quarterly) (bps)				
Panel A: alternate measures of ***distance***										
$distance_{t-1}$										
mov avg.	286***					247**				
	(2.98)					(2.43)				
no div		339***					277**			
		(2.96)					(2.60)			
non-zero			581***					407***		
			(5.37)					(4.67)		
vw herd				357***					262**	
				(3.14)					(2.63)	
FF48					340***					317***
					(4.38)					(4.62)
Controls	√	√	√	√	√	√	√	√	√	√
Observations	12996	16283	16283	16283	16277	12996	16283	16283	16283	16277
# groups	60	62	62	62	62	60	62	62	62	62
Panel B: alternate measures of ***direction***										
$direction_{t-1}$										
mov avg.	-228***					-220***				
	(-3.16)					(-4.38)				
no div		-94**					-93***			
		(-2.60)					(-3.80)			
non-zero			-50***					-40.43***		
			(-3.62)					(-2.86)		
vw herd				-92***					-78.15***	
				(-3.62)					(-4.64)	
FF48					-9					14.04
					(-1.46)					(1.24)
Controls	√	√	√	√	√	√	√	√	√	√
Observations	12996	16283	16283	16283	16277	12996	16283	16283	16283	16277
# groups	60	62	62	62	62	60	62	62	62	62

Table 3.9
Descriptors of Leaders and Followers

This table reports coefficients from a Pooled OLS specification regressing measures of leading and following on lagged manager and fund characteristics. The first three columns use $leader_t$ as the dependent variable and the last three columns use $follower_t$. All regressions include year and benchmark effects. Standard errors are clustered at the fund.

	$leader_t$			$follower_t$		
	[1]	[2]	[3]	[4]	[5]	[6]
$leader_{t-1}$		0.003				
		(0.21)				
$leader_{t-4}$			0.000			
			(0.02)			
$follower_{t-1}$					0.038***	
					(2.67)	
$follower_{t-4}$						0.042**
						(2.45)
mgr tenure_{t-1}	-0.0003	-0.0004	-0.0006	0.0001	-4.8e-05	0.0004
	(-0.70)	(-0.90)	(-1.22)	(0.34)	(-0.10)	(0.62)
fund age_{t-1}	0.0002	0.000	0.0005	-0.0002	-0.0003	-0.0003
	(0.67)	(0.06)	(1.23)	(-0.85)	(-0.78)	(-0.75)
perf pctl_{t-1}	0.0053**	0.0038	0.0012	-0.0079***	-0.0102***	-0.0047
	(2.20)	(1.44)	(0.36)	(-3.30)	(-3.43)	(-1.32)
$\text{flows/assets}_{t-1}$	0.0085*	0.0092	0.0113	0.0088*	0.0158***	0.0096
	(1.66)	(1.63)	(1.59)	(1.84)	(3.01)	(1.39)
turnover_{t-1}	0.0017*	0.002*	0.0004	-0.0045***	-0.0044***	-0.007***
	(1.66)	(1.69)	(0.25)	(-4.46)	(-3.49)	(-4.23)
expenses_{t-1}	-0.0001	-0.0001	0.002	-0.0007	-0.0022	0.003
	(-0.07)	(-0.06)	(0.72)	(-0.34)	(-0.92)	(0.87)
$\log(\text{tna})_{t-1}$	1.7e-07	2.8e-07	-2.2e-07	-2.2e-07	-2.9e-07	-5.7e-07*
	(0.44)	(0.68)	(-0.59)	(-0.90)	(-1.35)	(-1.74)
Observations	10924	9143	5212	13231	9091	5887
R-squared	0.13	0.13	0.09	0.15	0.14	0.13

Table 3.10
Performance, Leaders and Followers

This table reports results from a Fama-MacBeth specification regressing the fund's alpha on lagged measures of *leader* and *follower* behavior and controls. For each fund and quarter I estimate factor loadings using the previous 5 years of monthly returns from CRSP (at least 2 years required) and use these to compute the fund's 4-factor Carhart alpha (in bps). Control variables are fund age, expenses, turnover, fund flows, and log assets. Standard errors use Newey-West correction with maximum lags.

	Carhart α_t (quarterly) (bps)		
$leader_{t-1}$	202***		
	(7.76)		
$leader_{t-2}$		-17	
		(-0.74)	
$follower_{t-1}$			-128***
			(-3.25)
fund age$_{t-1}$	1.04	1.56***	0.86
	(1.59)	(3.11)	(1.48)
expenses$_{t-1}$	-5.95	-20.61***	-11.3
	(-0.66)	(-2.94)	(-1.26)
flows/assets$_{t-1}$	4.3	32	84.8***
	(0.11)	(1.61)	(3.03)
log(tna)$_{t-1}$	-0.001***	-0.004***	-0.003***
	(-2.34)	(-3.89)	(-3.22)
turnover$_{t-1}$	-10	6.3	-8.4
	(-1.60)	(1.07)	(-1.54)
Constant	-11.7	-12.8	-5
	(-1.00)	(-1.35)	(-0.52)
Observations	10679	10683	12929

Chapter 4

An Empirical Examination of the Impact of Herd Behavior on Stock Liquidity*

4.1 Introduction and Literature Review

A stock's liquidity and the risks that may arise from potential illiquidity are important factors for many investors in their investment decisions. Liquidity has been shown to not only affect stock returns, but to also covary strongly across stocks, i.e. commonality in liquidity.[1] This commonality in liquidity can arise from both supply-side and demand-side sources. While studies have found support for supply-side sources (e.g., Coughenour and Saad, 2004; Comerton-Forde, Hendershott, Jones, Moulton and Seasholes, 2010), other studies indicate that supply-side sources cannot explain all of the observed commonality in liquidity (e.g., Brockman and Chung, 2002; Bauer, 2004).[2] In this paper we propose that mutual funds should be large contributors to the demand-side source of commonality in liquidity, the rationale being that they are large investors with similar holdings and trading patterns who

*This chapter is based on research conducted in "Commonality in Liquidity: A Demand-side Explanation" co-authored with Stefan Ruenzi and Laura Starks.

[1]See, for example, Amihud and Mendelson (1986) and (1989), Brennan and Subrahmanyam (1996) Brennan, Chordia and Subrahmanyam (1998), Jacoby, Fowler and Gottesman (2000),Amihud (2002), Jones (2002), Chordia, Huh and Subrahmanyam (2007), and Hasbrouck (2009) regarding liquidity and returns and Chordia, Roll and Subrahmanyam (2000), Hasbrouck and Seppi (2001), Huberman and Halka (2001), Eckbo and Norli (2002) Brockman, Chung and Perignon (2007), Karolyi, Lee and vanDijk (2008) regarding commonality in liquidity.

[2]These papers find strong commonality in liquidity in pure limit order markets, while the explanation suggested in Coughenor and Saad (2004) is based on common market makers.

are often hit with the same liquidity shocks.

The intuition for our argument is as follows: The trades of certain groups of investors may exhibit similarity in both direction and timing. If a group of investors are subject to similar liquidity shocks or changes in their information set, the trades of these investors will likely be in the same direction (within a given stock) and occur with similar timing. If these investors hold a large group of stocks, then the stocks comprising their holdings are likely to experience large trade imbalances at the same points in time. It follows that stocks held to a large extent by a group of investors that tend to trade in the same direction and at the same time should be characterized by strong comovements in their liquidity.

Mutual funds are a prime example of an investor group that could give rise to such an effect. Mutual funds typically hold large, well-diversified portfolios and regularly face liquidity shocks in the form of positive or negative net-flows. The net-flows that mutual funds experience are typically highly correlated across funds, i.e., if one fund faces outflows (inflows), many others face outflows (inflows) at the same time. Furthermore, previous research has provided evidence of correlated trading by mutual funds as well as other institutional investors.[3] Consequently, we hypothesize that stocks with high mutual fund ownership should exhibit strong commonality in liquidity.

We test this basic hypothesis using an approach similar to that employed by Coughenour and Saad (2004) in their examination of the role of market makers in explaining commonality. Using data on mutual fund ownership and measures of stock liquidity for NYSE and

[3]See, for example, Kraus and Stoll (1972), Lakonishok, Shleifer and Vishny (1992), Grinblatt, Titman and Wermers (1995), Sias and Starks (1997), Wermers (1999), Sias (2004), Coval and Stafford (2007), Greenwood and Thesmar (2009), Anton and Polk (2010).

AMEX stocks over the 1980 to 2008 period, we estimate the covariance between a stock's liquidity and the liquidity of a portfolio of stocks with high mutual fund ownership, where we define liquidity by the Amihud (2002) measure of daily stock liquidity.[4] For the sake of brevity we label the regression coefficient on the high mutual fund ownership portfolio, HI, the mutual fund liquidity beta.

Our hypothesis implies a positive relation between HI and mutual fund ownership. To test this hypothesis, in each quarterly cross section we relate the stock's commonality of liquidity with the degree to which the stock is owned by mutual funds. We find that the liquidity of stocks with high mutual fund ownership covaries about twice as strongly with the liquidity of other high mutual fund ownership stocks than with the liquidity of stocks with low mutual fund ownership.

An alternative explanation for our findings is that mutual funds hold stocks with specific characteristics that explain commonality. That is, our results could be driven by individual stock characteristics such as firm size or level of liquidity that might jointly determine systematic liquidity and mutual fund ownership.[5] To test this alternative hypothesis, we conduct several refinements of our analysis. We examine the relationship between mutual fund ownership and the mutual fund liquidity beta within size and liquidity level quartiles. The positive relationship between mutual fund ownership and the mutual fund liquidity

[4]We control for market-wide commonality in liquidity when estimating the covariance by including the liquidity of the market portfolio in the time series regression. Coughenour and Saad (2004), in their analysis of the impact of common market makers on commonality, use the liquidity of a portfolio of shares that have the same market maker instead of the liquidity of a portfolio of high mutual fund ownership stocks as explanatory variable.

[5]See, for example, Del Guercio, 1996; Falkenstein, 1996; Gompers and Metrick, 2001; Bennett, Sias and Starks, 2003; Massa and Phalippou, 2005.

beta is strongest among large and liquid stocks, which tend to be the stocks most favored by mutual funds. However, the result also generally holds within all subsets except for the very smallest or most illiquid stocks, which is not surprising because mutual funds typically are not the dominant holders (or traders) of these types of stocks. Further, we also find the positive relation between mutual fund ownership and the mutual fund liquidity beta to continue to hold in a multivariate setting while controlling for the effects of a set of individual stock characteristics and even after including firm-fixed effects.

If the impact of ownership on commonality is driven by the trading activity of mutual funds, as we hypothesize, then one would expect the ownership-commonality relationship to be stronger under conditions in which ownership is a better proxy for correlated trading. To examine this, we consider the following two types of mutual fund trading: voluntary trading (often associated with information-based investment strategies) and involuntary trading (typically caused by liquidity shocks from fund flows).

A mutual fund's level of voluntary trading is reflected in the fund's turnover ratio after controlling for the fund's flow-induced trading. If a high proportion of the mutual funds' voluntary trading is due to correlations in information-based trading across funds, then we would expect a relation between the level of such trading and commonality in liquidity. Consistent with this hypothesis, we find that mutual fund liquidity betas are greater when stocks are owned by mutual funds with high turnover ratios than for stocks that are owned by mutual funds that do not trade a lot.

Involuntary or forced trading will be observed when mutual funds experience large inflows or outflows. This creates buying or selling pressure for those shares typically owned and traded by mutual funds (Coval and Stafford, 2007, Ben-Rephael, Kandel, and Wohl, 2009, and Khan, Kogan, and Serafeim, 2009). Furthermore, one would expect a difference between the effects of inflows and outflows as funds can accumulate cash before they have to trade based on inflows, but outflows can force the fund to eventually trade in order to meet redemptions (e.g., Edelen and Warther, 2001). We find strong evidence that suggests flow-driven liquidity shocks are an important driver of the effects of the mutual fund ownership results that we document. The impact of mutual fund ownership on a firm's mutual fund liquidity beta is about 50% greater in quarters with high absolute aggregate flows as compared to quarters with low absolute aggregate flows. The effect is particularly pronounced for negative flow quarters; the impact of ownership on commonality is roughly 75% stronger in quarters with highly negative net flows. This evidence supports the hypothesis that liquidity shocks that mutual funds face propagate through to the commonality in liquidity among the stocks they hold. These results also support the notion that liquidity demand of mutual funds contributes to commonality in liquidity.

Finally, in addition to using the level of ownership as a proxy for the likelihood of correlated trading we use the change in mutual fund ownership obtained from quarterly SEC filings. Consistent with our hypothesis, we find a strong positive relation between changes in a stock's aggregate mutual fund ownership and its mutual fund liquidity beta.

Our results are stable over time, hold over different subsamples, and are not driven by return or volatility comovements among stocks with high mutual fund ownership. Overall, our results suggest an important role for mutual fund ownership and eventually liquidity demand in explaining commonality in liquidity across stocks.

Our paper contributes to several main lines of research. It contributes to the broad empirical literature on liquidity in common stocks. A number of papers have documented the impact of liquidity on expected returns.[6] More recently, several studies document the existence of commonality in liquidity, in the U.S. as well as internationally.[7] Further the relevance of commonality for asset pricing is highlighted in both theoretical and empirical work.[8] The literature focusing on commonality in liquidity has focused on the supply side provision of liquidity. Coughenour and Saar (2004) show that commonality in liquidity can arise from the same NYSE specialist providing liquidity for many stocks. Consistent with this idea, Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes (2010) provide evidence that the aggregate inventory of all NYSE specialists is an important determinant of aggregate market liquidity. We contribute to this strand of the literature by showing the role of mutual funds in explaining commonality via the demand side.

The importance of the demand side of liquidity in explaining liquidity levels is provided by Chordia, Roll, and Subrahmanyam (2002) who find that aggregate order imbalance - which is a measure for liquidity demand - reduces liquidity. However, their focus is on liquidity levels, while our contribution is to show that liquidity demand has an impact on commonality of liquidity. While generally focusing on liquidity supply, Hameed, Kang, and Viswanathan (2010) also analyze the impact of correlated liquidity demand: consistent with our results, they find that comovements in stock-level order imbalance measures help to explain commonality. The impact of liquidity demanding trades on movements in market prices is also examined in Hendershott and Seasholes (2009). We add to this literature by

[6]See, for example, Amihud and Mendelson (1986), Brennan, Chordia and Subrahmanyam (1998), Jacoby, Fowler and Gottesman (2000), Jones (2002), Amihud (2002), Chordia, Huh and Subrahmanyam (2007), and Hasbrouck (2009).

[7]See, for example, Chordia, Roll and Subrahmanyam, (2000), Hasbrouck and Seppi (2001), Brockman, Chung and Perignon (2007), and Karolyi, Lee and vanDijk (2008).

[8]See, for example, Pastor and Stambaugh (2003), Acharya and Pedersen (2005), Sadka (2006), Korajczyk and Sadka (2008), and Lee (2010).

identifying a primary source of the comovements.

Our findings also contribute to the literature on the influence of investors, particularly institutional investors, on stock returns.[9] With regard to liquidity effects, Massa (2004) and Massa Phalippou (2005) examine the relation between institutional investor ownership and the level of stock liquidity. Kamara, Lou and Sadka (2008) examine the impact of changing aggregate levels of institutional ownership on commonality in returns, finding that commonality increases over time. Consistent with our results, they argue that this is driven by the increasing importance of institutional investors over time. Further in terms of the impact of investors' correlated trading on returns, Greenwood (2009) shows that common trading patterns of index investors can give rise to substantial excess comovement of stock returns. Pirinsky and Wang (2004) and Kumar and Lee (2006) find that correlated trading among institutional and retail investors, respectively, gives rise to return comovement.[10]

More closely related to our paper are Greenwood and Thesmar (2009) and Anton and Polk (2010). Greenwood and Thesmar also use mutual fund ownership and mutual fund flows to get a proxy for correlated trading. Examining the 1990 to 2008 period they show that stocks owned by mutual funds with correlated inflows exhibit larger return comovements. Anton and Polk provide evidence that common covariation in stocks is associated with common ownership by mutual funds. We contribute to their findings by showing the channels through which institutional investors can give rise to commonality in returns. In addition, none of these papers investigates the link between correlated trading and comovement in

[9]See, for example, Sias and Starks, 1997; Gompers and Metricks, 2001; Sias, Starks and Titman, 2006.

[10]Evidence suggesting that investor clienteles might lead to return comovement is also provided in Barberis, Shleifer, and Wurgler (2005), Pirinsky and Wang (2006), and Green and Hwang (2009).

liquidity.

The remainder of this paper is organized as follows. In Section 4.2 we describe our data and the construction of our main variables. Our empirical analysis regarding commonality in liquidity and mutual fund ownership is presented in Section 4.3 and in Section 4.4 we consider proxies for mutual fund trading. We provide results from robustness tests in Section 4.5 and our conclusions in Section 4.6.

4.2 Data and Variable Construction

Our initial sample is based on mutual fund holdings from the CDA/Spectrum database over the 1980-2008 period.. We match the holdings of these mutual funds to other fund variables in the CRSP mutual fund database using MFLinks. We also match these data to characteristics of the underlying stocks obtained from the CRSP stock database.

4.2.1 Variable Definitions

Ideally we would be able to directly observe mutual fund trades in order to measure each stock's degree of correlated mutual fund trading through time. Because we have quarterly snapshots of mutual fund ownership rather than trades, we create a stock-level proxy for the likelihood of correlated trading based on the percentage of shares outstanding held by mutual funds. Specifically, for each stock we construct a quarterly measure of aggregate mutual fund ownership.[11] The fraction of ownership $mfown_{i,t}$, in stock i owned by J mutual

[11] To obtain quarterly stock level measures of aggregate mutual fund ownership using March, June, September, and December as quarter end dates we carry forward each fund's quarterly holdings for two months. Then, following the literature, we carry holdings forward an additional quarter if the fund appears to have missed a report date (see, e.g., Frazzini and Lamont, 2008). This is done for a maximum of a 6 month

funds at the end of quarter t, is

$$mfown_{i,t} = \frac{\sum_{j=1}^{J} sharesowned_{i,j,t}}{shrout_{i,t}},$$

where $sharesowned_{i,j,t}$ is the number of shares in stock i owned by mutual fund j at quarter t and $shrout_{i,t}$ is the total number of shares outstanding. We use this variable as a proxy for the likelihood that, for a given stock, large order imbalances will occur.

In our later analysis, we also consider a further measure of mutual fund trading and use a turnover-weighted version of $mfown_{i,t}$. When summing ownership across funds within a stock, we weight ownership by turnover,

$$twmfown_{i,t} = \frac{\sum_{j=1}^{J} turnover_{j,t} * sharesowned_{i,j,t}}{shrout_{i,t}},$$

where $turnover_{j,t}$ equals the turnover as reported in CRSP for fund J during quarter t.

We measure liquidity using the Amihud (2002) measure of daily stock illiquidity, which equals the absolute value of return for stock i on day d divided by the dollar volume of trading for stock i on day d. The Amihud measure is ideal for our research because it is based on widely available data and can be calculated for a large number of stocks at a daily frequency. Evidence also supports the use of the Amihud measure as a reliable proxy for a stock's liquidity with strong correlations between it and alternative liquidity measures based on intraday microstructure measures (e.g., Hasbrouck (2005) and Koraczyk and Sadka (2008).) More recently Goyenko, Holden, and Trzcinka (2009) show that the Amihud (2002) measure is a good proxy for price impact.

gap in report dates. Holdings are adjusted for splits that occur between the reporting and filing dates. We set holdings equal to zero if the report date is subsequent to the file date, if CRSP reports zero shares outstanding, or if the total mutual fund ownership exceeds the shares outstanding.

The Amihud (2002) measure comes into our analysis in two ways. First, we use the quarterly average of the daily Amihud illiquidity measure as a control variable in many of the regressions to take into account the potential impact of the level of stock liquidity. Second, for our primary variable we employ the change in the Amihud (2002) illiquidity measure. Specifically, we compute the change in the daily measure of stock illiquidity using volume and return data from CRSP,

$$\Delta illiq_{i,d} = ln\left(\frac{illiq_{i,d}}{illiq_{i,d-1}}\right) = ln\left(\frac{\frac{|r_{i,d}|}{|dvol_{i,d}|}}{\frac{|r_{i,d-1}|}{|dvol_{i,d-1}|}}\right),$$

where r_i, d is the return on stock i for day d and $dvol_{i,d}$ is the dollar volume for stock i on day d.[12] We calculate the daily change in stock illiquidity for all common stocks on the NYSE and AMEX that are not penny stocks (i.e., price is above $2 per share), that trade on day d and d-1, and that have at least 40 return observations in a quarter. To prevent outliers from affecting our analysis, we eliminate the top and bottom 1% of observations of our measure.

4.2.2 Summary Statistics

Table 4.1 reports statistics on the sample stocks' market value, illiquidity measure, mutual fund ownership, and mutual fund ownership weighted by fund turnover. The table also reports statistics for aggregate quarterly mutual fund flows. Panel A shows the statistics across all stocks and quarters for which we have data. The final sample consists of 120,413 stock-quarters with both mutual fund ownership data and sufficient data to calculate liquidity

[12]By taking the difference of the logs of Amihud's illiquidity measure we follow Kamara, Lou, and Sadka (2008). This is done to reduce effects of non-stationarity. However, in light of concerns of over-differencing, we also replicate the main results using the difference in Amihud's illiquidity measure from its five day moving average (see Section 4.5).

betas. Using the turnover-weighted mutual fund ownership reduces the sample to 66,598 stock-quarters because turnover data is only available beginning in 1999. The median firm has $897 million in market equity and 10% of its shares are owned by mutual funds. The mean turnover-weighted mutual fund ownership is slightly smaller than un-weighted mutual fund ownership, reflecting a typical annual fund turnover ratio of less than one (in our sample the average fund turnover is 0.83). In the last row we report summary statistics on aggregate quarterly net-flows into or out of the equity mutual fund industry. Over our sample period (1980-2008) mutual funds generally experience inflows, however aggregate flows are negative in 17 of the quarters with the largest aggregate quarterly outflow equaling 3.05% of the NYSE and AMEX market capitalization, compared to the largest aggregate quarterly inflow of 2.83%.

Panel B of Table 4.1 shows the summary statistics by quartile of mutual fund ownership. In each quarter we rank stocks by $mfown$ and report means, standard deviations, and medians of the selected variables. Typical stock size is about $3 billion in the lowest and highest quartiles of mfown compared to $7 and $4 billion for the second and third quartiles respectively. There is, however, a monotonic relationship between mutual fund ownership and average liquidity; moving from the lowest to highest quartile of $mfown$, the $illiq(avg)$ drops from 0.19 to 0.04.

4.3 Commonality in Liquidity and Mutual Fund Ownership

In order to examine the extent to which mutual fund ownership determines comovement in liquidity, we follow an approach similar to that in Coughenour and Saad (2004). In the first step, we estimate how individual stock liquidity co-moves with the liquidity of a portfolio of high mutual fund ownership stocks after controlling for comovement with market liquidity and additional variables (Section 4.3). In the second step we investigate whether comovement between individual stocks and the high $mfown$ portfolio is stronger among firms with high mutual fund ownership (Section 4.3).

4.3.1 Estimating Liquidity Covariances

We first estimate for each firm-quarter the covariance between the daily changes in a stock's illiquidity and changes in the illiquidity of a portfolio of stocks with high mutual fund ownership. We control for the widely documented comovement in individual illiquidity with market illiquidity (Chordia, Roll and Subrahmanyam, 2000). Thus, for each trading day in the quarter we compute changes in the value-weighted illiquidity of two portfolios: a market portfolio containing all stocks and a high mutual fund ownership portfolio comprised of the stocks in the top quartile of mutual fund ownership as ranked at the end of the previous quarter.[13]

For each firm, we run quarterly time series regressions of the firm's daily change in illiquidity, $\Delta illiq_{i,t}$, on changes in the high mutual fund ownership portfolios' illiquidity, $\Delta illiq_{mfown,t}$, and changes in the market illiquidity, $\Delta illiq_{mkt,t}$, as well as control variables:

$$\Delta illiq_{i,t} = \alpha + \beta_{HI}\Delta illiq_{mfown,t} + \beta_{mkt}\Delta illiq_{mkt,t} + \delta controls + \varepsilon_{i,t}, \tag{4.1}$$

We focus on changes, or to be precise changes in logs, because we want to investigate the similarity in movements in liquidity. Furthermore, this approach helps to avoid

[13]Results using equal-weighted portfolios are very similar (see Section 4.5).

econometric problems due to the potential nonstationarity of the liquidity measure. For each regression, the firm of interest is removed from the market portfolio as well as the high mutual fund ownership portfolio (when applicable). We follow the approach taken by Chordia, Roll, and Subrahmanyam (2000) and include lead, lag and contemporaneous market returns, contemporaneous firm return squared, and lead and lag changes in the two portfolio illiquidity measures. The latter controls are designed to capture lagged adjustments in liquidiity, while the market returns are included to control for possible correlations between returns and our illiquidity measure. The squared stock returns are included to capture volatility which might be related to liquidity. We require a minimum of 40 observations for each firm-quarter.[14] We show later in robustness tests (Section 4.5) that this particular specification of the first stage time series regressions is not crucial to our main results.

Table 4.2 presents sample statistics on the market and high mutual fund ownership portfolios used in the time series regressions as well as coefficients of interest from the regressions. In Panel A we report summaries for representative quarters, one each from the beginning (1980), the middle (1995) and the end (2008) of our sample. In Panel B we summarize by 5 year periods as well as the full sample.

The left-hand side of each panel reports the average of the mutual fund liquidity beta coefficients across all firms in that quarter, the percentage of beta coefficients that are positive and the percentage that are significant as well as a t-statistic on the sample of beta coefficients in that quarter. The table also reports the number of stocks in the portfolio and the average firm size and illiquidity.

[14]Results are very similar if instead of requiring a minimum of 40 observations we require a minimum of 30 or 50 observations.

Relatively few of the beta estimates are significantly different from zero at the 5%-level based on two-sided t-tests. This is likely due to the large noise in the firm level regressions, which are conducted on a quarterly basis.[15] While few of the individual quarterly estimates are statistically significant, the mean of the distribution of estimates is different from zero with a high degree of significance as indicated by the t-statistic on the sample of estimates. The right-hand side of the table summarizes the same variables for the market liquidity beta coefficients. Overall, the positive average and the similar magnitude of the two beta coefficients, β_{HI} and β_{mkt}, clearly shows that individual stock liquidity on average co-moves positively with both the liquidity of the market portfolio as well as the liquidity of a high mutual fund ownership portfolio. However, in the next section we test our main hypothesis: that β_{HI} is higher among stocks with high mutual fund ownership.

The bottom panel summarizes the time series regression output by 5 year periods. We calculate summary variables and t-statistics for each quarter as above, and in this panel we report averages of these quarterly summary variables. For example, in the 1980-1985 period the typical quarter has a mean β_{HI} equal to 0.26 and the average t-statistic on each quarter's sample of estimates is 5.10.

The average size of firms in the high mutual fund ownership portfolio is smaller than the average size of the firms in the market portfolio. Average mutual fund ownership over the entire sample of stocks is increasing through time. The average mutual fund ownership in a stock is 4% in 1980 and this number increases to 24% in the third quarter of 2008.

[15]In unreported tests, using the full available time series for each stock we find that 71% of the market liquidity betas and 77% of mutual fund liquidity betas are positive, with 24% and 28% significantly different from zero at the 5% level, respectively.

Among the stocks in the top quartile of mutual fund ownership, average ownership increases from 9% in 1980 to 37% in 2008. Stocks were less liquid in the 1980's relative to the later period. This finding is consistent with the results in Jones (2002). The decrease in illiquidity is most pronounced among the stocks in the highest quartile of mutual fund ownership. The average illiquidity among the stocks in this portfolio is lower than the average illiquidity of the stocks in the market portfolio in all quarters. This result shows that mutual funds prefer liquid stocks, which is also similar to results from earlier studies (e.g., Falkenstein, 1996).

4.3.2 Mutual Fund Ownership and Commonality

Our central hypothesis is that the liquidity of stocks with high levels of mutual fund ownership will covary strongly with other stocks also owned to a high degree by mutual funds. Table 4.3 provides results from a first set of tests of our central hypothesis using one dimensional and dependent sorts based on quarterly rankings of mutual fund ownership. In this and all future tests, $\beta + HI$ and β_{mkt} are estimated over quarter t, while mutual fund ownership is measured at the end of quarter $t-1$.

Panel A shows that the average β_{HI} is monotonically increasing in mutual fund ownership as predicted by the hypothesis. The lowest ownership quartile has an average Panel A shows that the average β_{HI} is monotonically increasing in mutual fund ownership as predicted by the hypothesis. The lowest ownership quartile has an average HI of 0.20 compared to 0.40 for the highest quartile. The difference is economically and statistically significant, providing evidence that the liquidity of stocks owned to a high degree by mutual funds strongly covary together. These findings provide first evidence for our central hypothesis.

The results for β_{HI} can be contrasted with those for β_{mkt} reported on the right hand side of Panel A. There is no significant difference between the comovement of stocks' liquidity with the overall market liquidity in the highest and lowest mutual fund ownership quartiles.

We also report averages for β_{HI} and β_{mkt} from sorts based on firm size and liquidity. For β_{HI}, the difference between the top and bottom quartiles is statistically significant in both cases. Large stocks have a significantly higher average β_{HI} of 0.29 compared to 0.23 among the smallest quartile. However, the relationship is non-monotonic. We find a similar non-monotonic relationship between average illiquidity and β_{HI}. There are also strongly significant differences between the comovement of a stock's liquidity with the market liquidity in the highest and lowest size and illiquidity, respectively, quartiles. Our results show that large and liquid stocks co-move more heavily with both market as well high mutual fund ownership portfolio liquidity as compared to small and illiquid stocks.

The results presented thus far are univariate in nature. However mutual funds do not randomly select stocks but have preferences for certain characteristics. Most importantly, in aggregate they prefer large and liquid stocks (see, e.g., Del Guercio, 1996; Falkenstein, 1996). Our previous results suggest that these characteristics are also related to β_{HI}. Thus, in Panel B of Table 4.4 we provide the results on the average liquidity betas for double sorts based on these variables and mutual fund ownership. The sorting is first done on size or illiquidity and then on mutual fund ownership. The results show that the positive relation between HI and mutual fund ownership is robust to subsets by firm size and illiquidity. In all cases the average β_{HI} is increasing in mutual fund ownership although the effect is insignificant among the most illiquid stocks. The latter are the stocks that are least held by mutual funds, which we expect would not be much affected by correlated mutual fund stock trading.

In a second test of our central hypothesis we control for stock characteristics in a multivariate regression. We regress β_{HI} against the previous quarter's mutual fund ownership, controlling for firm size and average illiquidity. We include time dummies and cluster the standard errors at the firm level. This accounts for time series and cross sectional dependence as long as the time effect is fixed (Petersen (2009)). The specification is

$$\beta_{HI} = a + b_1 mfown_{i,t-1} + b_2 ln(size_{i,t-1}) + b_3 illiq(avg)_{i,t-1} + timedummmies + \varepsilon_{i,t}. \quad (4.2)$$

Our main hypothesis predicts b_1¿0. We do not have clear theoretical predictions on b_2 or b_3. However, given the results from Table 4.3, one might expect a positive relation between β_{HI} and firm size and a negative relation with illiquidity. The results of this regression are presented in Panel A of Table 4.4. The first column of the table shows the results for the full sample for the regression of β_{HI} against mutual fund ownership and time dummies only. We affirm that stocks with high mutual fund ownership exhibit strong comovement, evidenced by the significant coefficient estimate of 0.896. As this regression includes time fixed effects, the higher β_{HI} should not be caused by a possible common time trend in mutual fund ownership levels and liquidity comovements.

In Model (2) we control for the stock's size and average liquidity. Again the coefficient on mutual fund ownership is positive and highly significant, and is similar in magnitude to the coefficient estimated in the absence of controls. The result is also economically significant - a one standard deviation increase (0.10) in mutual fund ownership is associated with a 0.08 increase in HI, which equates to a 27% increase from its mean.

4.3.3 Potential Alternative Explanations and Specifications

Another possible explanation for our results is that mutual fund managers have preferences for a stock characteristic (other than size and liquidity) that is correlated with β_{HI}. Although it is not clear what the source of the unobserved heterogeneity and correlation might be, in Model (3) we include firm fixed effects to address this concern. We continue to include time dummies and cluster standard errors at the firm level. The results show that time invariant unobservable heterogeneity is not driving our results.

The last two models in Table 4.4 use corrections for different assumptions on the structure of the error term. Model (4) employs standard errors with two dimensional clustering, and Model (5) uses a Fama-MacBeth specification. In both alternative models we find a positive relationship between the mutual fund liquidity beta and mutual fund ownership that is both economically and statistically significant.

We have no direct prediction on the functional form of the relationship between ownership and commonality, and so for further robustness we repeat our tests using an indicator variable for high mutual fund ownership rather than a continuous variable. We replace $mfown_{i,t-1}$ in (2) by $mfown(dummy)_{i,t-1}$, which is equal to one if mutual fund ownership is in the top quartile in quarter $t-1$, and zero otherwise. These results are reported in Panel B of Table 4.4. The use of this variable provides a natural economic interpretation. From column 2 in Panel B, stocks in the highest mutual fund ownership quartile have a β_{HI} in the next quarter that is 0.12 higher than those outside the top quartile. This is a large economic effect given the unconditional mean β_{HI} of 0.31. The coefficient on this dummy variable is positive and statistically significant in all other specifications as well.

The sorts in Table 4.3 indicate a possible non-linear relation between β_{HI} and firm size or illiquidity. Thus, we rerun our primary multivariate specification (quarter fixed effects and firm clusters) for samples divided by size quartiles, additionally controlling for size and liquidity within each subsample. We also conduct this test for subsamples divided by liquidity, time (5 year subperiods), and whether the quarter has an up or down market return. Table 4.5 reports these results again for a linear impact of $mfown$ (Panels A and B) as well as for the impact of the high mutual fund ownership dummy (Panels C and D).

In Panels A and C, the first four columns split the sample into size quartiles (ranked quarterly) and show that a significantly positive relation between β_{HI} and mutual fund ownership exists in all but one of the subsamples, the quartile of stocks with the smallest market capitalization. The next four columns report the results from the sample divided into liquidity quartiles and show a significantly positive relationship between β_{HI} and mutual fund ownership in all but the most illiquid stocks. This result is consistent with our results using dependent sorts in Panel B of Table 4.3.

When we divide our sample into approximate 5-year subsamples from 1980 to 2008 (with the last subperiod containing almost 8 years), we find that the effect exists in all subperiods, but the magnitude of the coefficient for the relation between β_{HI} and mutual fund ownership varies over time.

Motivated by results of magnified liquidity effects in down markets in Chordia, Roll and Subrahmanyam (2002) and Hameed, Kang, and Viswanathan (2010), we also look at subsamples of up as well as down market quarters. We find a strong effect in both cases. The coefficient on $mfown$ is larger in quarters with negative market returns. However, the difference between the coefficients in the up versus down market subsamples is not significant.

This shows that although previous research documents higher commonality in liquidity in down markets, up versus down markets have no impact on the role of mutual fund ownership in explaining liquidity. Rather, results are fairly stable across market regimes.[16]

Overall, these results provide solid evidence that the liquidity of stocks with high mutual fund ownership strongly co-move. The effect is robust to various assumptions regarding unobserved heterogeneity, independence of observations, and functional form, as well as a variety of subsamples.

4.4 Commonality in Liquidity and Mutual Fund Trading

In the previous section we provide evidence that commonality in a stock's liquidity is strongly associated with the level of mutual fund ownership in the stock. We claim that this relationship exists because mutual fund ownership proxies for the likelihood that trading will be correlated. That is, it is not the level of ownership that matters per se, but the extent to which it reflects future correlated trading. In the following section we test alternative proxies for the probability of future correlated trading.

In the absence of directly observing trades, an ideal proxy would reflect two probabilities, i) the likelihood that a stock is traded and ii) conditional on being traded the likelihood that the trades are in the same direction. We refine $mfown_{i,t}$ in three ways to capture the likelihood of future correlated trading; a measure that reflects correlated voluntary trading, one that reflects correlated forced trading, and one that reflects overall correlated trading.

[16] In unreported results we examine differences between the levels of market-wide commonality in up and down markets and confirm the results of Chordia, Roll and Subrahmanyam (2002) and Hameed, Kang, and Viswanathan (2010) in our sample.

The first proxy allows for differential trading among mutual funds by incorporating the fund's turnover ratio into the ownership measure. That is, we treat ownership by high turnover funds as a better proxy for the likelihood of correlated trading than the same level of ownership by funds with low turnover. Because the turnover ratio as reported in CRSP is corrected for trading due to flows, it reflects voluntary trading. However, voluntary trading could reflect trading by mutual funds providing liquidity to other market participants as well as their information-based liquidity demanding trades (Da, Gao, and Jagannathan, 2008). While both cases could explain commonality, only the latter would be consistent with mutual funds demanding liquidity and eventually giving rise to commonality via this channel. Thus, to investigate whether the mutual fund demand side channel plays an important role in commonality of liquidity, we include a measure of future correlated trading designed to capture the effects of liquidity shocks to the fund itself due to inflows or outflows. Therefore our second refinement is to condition mutual fund ownership on aggregate fund flows. Because flows can lead to buying or selling pressure of mutual funds, i.e. liquidity demand, if commonality among mutual fund owned stocks is higher in periods of high absolute flows (and particularly in periods of high outflows), this is a clear indication that mutual funds have an impact on commonality via their liquidity demand.

Our final refinement is to use changes rather than levels of ownership. The change in ownership reflects actual trades in the same direction, thus capturing both the probability a stock is traded and the probability that trades are in the same direction. Therefore it should not be surprising that the change in ownership - at atomistic granularity - is the ideal measure. However data availability limits us to quarterly changes. Thus, using changes presents the tradeoff of measuring some fraction of trading with certainty, but also underestimating the amount of actual trading.

4.4.1 Mutual Fund Turnover

As a first approach to better capture the probability of correlated trading, we incorporate mutual funds' turnover ratios. When summing ownership across funds within a stock, we weight mutual fund ownership by the holding fund's turnover, turnover-weighted mutual fund ownership, $twmfown_{i,t}$ as defined in Section 4.2.

We expect to find that the turnover-weighted measure, to the extent that it is a better proxy for correlated demand in liquidity, is more strongly associated with high commonality in liquidity than an unconditional measure of mutual fund ownership.[17] One drawback of this refinement is data limitation because CRSP does not report fund turnover prior to 1999. The results are reported in Table 4.6. The first model includes twmfown only. For comparison, the second column repeats the evaluation of our baseline model using $mfown$ as the primary independent variable for the limited sample 1999 to 2008. It should be noted that the results for $mfown$ in this restricted time period are consistent with the results for the full sample period reported in Tables 4.4 and 4.5. The model reported in the third column includes both $twmfown$ and $mfown$.[18] The coefficient on the turnover-weighted mutual fund ownership variable is strongly significant in all three models irrespective of whether un-weighted mutual fund ownership is included.

[17]Importantly, this would not be case if there exists a negative relationship between correlated trading and fund turnover strong enough to outweigh the high levels of trading reflected by high fund turnover.

[18]The correlation between mfown and twmfown is 0.78, which might hint at multicollinearity in the model including both variables. However, the significant impact of twmfown we find as well as the relatively low variance inflation factors of 3.68 and 2.97 for mfown and twmfown, respectively, clearly indicate that this is not a concern here.

The summary statistics reported in Table 4.1 show sufficient similarity in the means and standard deviations of the weighted and unweighted mutual fund ownership measures, which suggests that we can roughly compare the coefficients of the two measures. Such a comparison shows that the coefficient for the turnover-weighted mutual fund ownership measure in column 3 is 1.152, which is clearly larger than the coefficient for the unweighted mutual fund ownership, which is 0.185 and not statistically distinguishable from zero. To provide a more precise comparison in the last three models of the table we use standardized independent variables. Again the results indicate that ownership by mutual funds with greater portfolio turnover is associated with higher commonality in liquidity than simply ownership by mutual funds in general. Further, Column 6 shows that a one standard deviation increase in $twmfown$ is associated with a 0.09 increase in β_{HI}. Thus, consistent with our hypothesis, stocks held by mutual funds that trade more frequently have stronger commonality in their liquidity.

Voluntary trading is often information-based trading. Thus, the strong impact of voluntary mutual fund trading on commonality suggests that the trading of individual mutual funds does not cancel out. This is consistent with the view that mutual funds tend to trade on the same information in the same direction, which eventually leads to correlated liquidity demand and thus commonality in liquidity.

An alternative story to explain these results is that voluntary trading is not information driven (and thus a sign of liquidity demand), but that mutual funds also act as liquidity suppliers in some cases. Thus, in the following section we focus on the impact of liquidity shocks mutual funds face themselves. This will allow us to isolate cases in which any potential effect most likely works via a demand-side channel.

4.4.2 Aggregate Fund Flows

In the previous section we investigate the relation between β_{HI} and a proxy for voluntary mutual fund trading. In this section we estimate the relation between β_{HI} and involuntary correlated trading. Thus, we infer differences in trading intensities using fund flows.[19] According to our hypothesis, the impact of mutual fund ownership should be greater in periods with high absolute flows. This effect should be particularly strong for outflows as suggested by the results of Coval and Stafford (2007). The reason why we expect a stronger impact of outflows is that inflows can be more easily spread across stocks, but fund outflows, if met through stock sales, must be met by selling the stocks currently held by the mutual funds.[20]

To examine the impact of flow levels, in each quarter we aggregate fund flows to compute a net dollar flow into or out of equity mutual funds. We then scale this amount by the dollar value of the total market at the beginning of the quarter. From the flow data we calculate four dummy variables; $negnetflow$ equals one if aggregate flows are negative, and zero otherwise, and $hiabsflow$ equals one if aggregate flows in a quarter are in the top or bottom 10% of all quarters, and zero otherwise. Net flows are signed, so the bottom (top) 10% is comprised of the largest net outflow (inflow) quarters. We also aggregate inflows and outflows separately in each quarter. Then we set $hioutflow$ ($hiinflow$) equal to

[19] Chordia, Roll, and Subrahmanyam (2009) find that fund flows can explain much of the increased turnover in equity markets over recent years. Furthermore, mutual funds tend to scale up their existing holdings if they face inflows of new money (Pollet and Wilson, 2008), i.e. inflows should lead to liquidity demand for those stocks with high previous mutual fund ownership.

[20] That high negative mutual fund flows lead to correlated liquidity demand is also suggested by the findings of Hameed, Kang, and Viswanathan (2010) who document a negative relation between commonality in order imbalances and aggregate net fund flows.

one for the top 25% of quarters measured by outflows (inflows) scaled by market cap, and zero otherwise. Each of these dummy variables is interacted with $mfown$ in the previously described regression specifications used in Table 4.4. We continue to use time dummies to pick up general increases or decreases in systematic liquidity during periods of extreme flows.

The results of these regressions are reported in Table 4.7. The results of Model 1 show that the impact of ownership on commonality is much stronger during periods of high positive or negative net flows. Specifically, the coefficient on $mfown$ is 0.765 in 80% of the quarters compared to 0.765 + 0.395 = 1.160 in the top and bottom 10% of flows (strong inflows and outflows). The results of the alternative specifications reported in Columns 2, 3 and 4 are similar. In column 2 the relation between β_{HI} and $mfown$ is 0.575 larger when the mutual fund industry experiences net outflows relative to the quarters with net inflows. This effect is highly significant both economically and statistically. Columns 3 and 4 show that we find similar results when aggregating inflows and outflows separately in each quarter. These results are consistent with the hypothesis that fund flows lead to correlated liquidity demand by mutual funds and that this effect is more pronounced for outflows. These results are also consistent with those of Coval and Stafford (2007) regarding mutual fund fire sales.

Columns 5 through 8 show the results from our base regression (2) within subsamples of quarters split by the level of aggregate funds flows. The strong relation between commonality in liquidity and mutual fund ownership holds in each of the subsamples. There is some evidence of a U-shaped relationship between the magnitude of liquidity commonality and aggregate net flows, as would be expected if mutual fund ownership has a larger impact during periods of extreme flows. However, consistent with results from the interactions in columns 2 through 4, this seems primarily driven by negative flow quarters. In Panel B of Table 4.7

we test specifically for a U-shaped conditional relationship. First, we run 114 quarterly cross sectional regressions based on model (2), regressing commonality on ownership and controls. Then we use the time series of coefficients on $mfown$ as the dependent variable in a regression with aggregate net flows and squared aggregate net flows as independent variables. We find that the impact of ownership on commonality is strongest in periods of high inflows and outflows as evidenced by the positive coefficient on aggregate flows squared, and that the effect of outflows dominates the effect of inflows, as evidenced by the negative coefficient on aggregate flows.

Overall, the findings from this section show that, in addition to voluntary information-based trading, flow induced liquidity demanding trades give rise to commonality in liquidity.

4.4.3 Changes in Mutual Fund Ownership

Finally we use actual changes in mutual fund ownership of individual stocks through the holdings data. Specifically we compute the absolute value of the change in $mfown_i$ from $t-1$ to t, and denote this variable $|\Delta mfown_{i,t}|$. The change in ownership reflects an amount of trading that we can be certain took place, and that these trades were in the same direction. We are limited by data availability to compute changes on a quarterly basis. Therefore while changes in ownership reflect with certainty some amount of correlated trading, an important drawback is that this captures only the lower bound.

We measure the change contemporaneously with the estimation of β_{HI} to determine whether higher sensitivity to aggregate mutual fund liquidity occurs in the same period as greater mutual fund trading, which would be consistent with correlated trading by mutual funds contributing to commonality in liquidity. We employ the following specification for this test:

$$\beta_{HI} = a + b_1|\Delta mfown_{i,t}| + b_2 ln(size_{i,t-1}) + b_3 illiq(avg)_{i,t-1} + timedummmies + \varepsilon_{i,t}. \quad (4.3)$$

A positive and significant b1 would support our hypothesis.

The results of this regression are provided in Table 4.8. We use the absolute value of the change in $mfown$ in the first model, and a dummy variable equal to one if the absolute change is in the top quartile that quarter, and zero otherwise, in the second model. In both cases the coefficient on the change measure is positive and significant at the 1% level, consistent with our hypothesis that mutual fund trading in a stock as reflected by changes in a stock's mutual fund ownership increases systematic liquidity.

Overall, the results of Tables 4.6, 4.7, and 4.8 clearly support our hypothesis that the relation between commonality in liquidity and mutual fund ownership is due to correlations in the trading by mutual funds.

4.5 Robustness Tests

Thus far, we have shown that the relationship between β_{HI} and $mfown$ is robust to different specifications regarding functional form and structure of the error term. We find additional support for our hypothesis through several refinements of our main variable of interest, turnover-weighted $mfown$, $mfown$ conditional on flows, and changes in $mfown$. In this section we address concerns arising from our first stage estimate of common liquidity, and in particular our use of the Amihud illiquidity ratio as the measure of liquidity. For example, the commonality that we document may be driven by common (absolute) returns, not necessarily common movements in the ratio of returns to volume. In this section we first demonstrate that our results are not driven by common returns or common volatility, and then show that our results are not specific to the structure or our first stage estimation.

We address a potential impact of common returns and common volatility in three ways. First, we add beta estimates between the firm return and the value-weighted return of the high mutual fund ownership portfolio (estimated contemporaneously with the liquidity beta) as an additional control variable in our base regression equation (2). We call this variable mutual fund return beta. Adding that variable controls for the impact of common information – that has a joint impact on the returns of the stocks with high mutual fund ownership - on the comovements in liquidity. Results are presented in the first column of Panel A in Table 4.9. Regarding the new control variable, we find a significantly positive impact of the mutual fund return beta on β_{HI}. This shows that common return effects (as a proxy for information affecting the returns of high mutual fund ownership stocks) also has an impact on commonality in liquidity among these stocks. More importantly, the positive impact of mutual fund ownership on β_{HI} still remains highly significant and is only slightly reduced after inclusion of the mutual fund return beta as compared to the results reported in Table 4.4. Second, to capture any potential non-linear relationship between β_{HI} and return comovements, we run our base regression (2) on subsamples based on mutual fund return beta quartiles. Results reported in columns 2 through 5 show that our main finding holds in all subsamples as indicated by a highly significant positive estimate for the impact of $mfown$ on β_{HI} in each case. Third, we modify the first stage regression (1) in order to capture the impact of a potential comovement between individual stock liquidity and the return of the portfolio of high mutual fund ownership stocks. Thus, we include the return of a portfolio of high mutual fund ownership stocks as additional control variable in (1). Results from equation (2) using the β_{HI} from this modified first stage model as dependent variable are presented in column 6 in Panel A of Table 4.9.[21] We still find a highly significant positive impact of $mfown$ on β_{HI}.

One may also be concerned that our results are driven by comovements in volatility among stocks with high mutual fund ownership which might be caused by joint changes in the riskiness of the funds owned by mutual funds. To address this we conduct the same battery of tests as above, but now replace the return by the return squared (for both the individual stock and the high mutual fund ownership portfolio), i.e. we use squared returns as volatility proxy. Results in Panel B of Table 4.9 show that our earlier results hold: the positive relationship between $mfown$ and β_{HI} is highly significant also after controlling for comovements in volatility (mutual fund return^2 beta; columns 7 through 11). Adding the squared return of the high mutual fund ownership portfolio in the first stage regression (to control for the impact of the comovement of individual liquidity and high mutual fund ownership portfolio volatility) does also not change the results obtained from the standard second stage regression (column 12).

Finally, we repeat our whole analysis using turnover instead of the Amihud illiquidity ratio as an alternative liquidity measure.[22] Results are presented in the first column of Table 4.10. There continues to be a strong positive relationship between ownership and commonality using the alternative liquidity proxy.

[21] We find similar results if we include market returns instead of or additionally in model (1).

[22] We use the Amihud measure in our main examination, because stock turnover is only a weak proxy for liquidity and is also mechanically related to our measure of turnover-weighted mutual fund ownership, because trading of mutual funds is directly linked to turnover on the stock level.

Overall, these findings show that our previous results are not driven by return or volatility comovements among stocks with high mutual fund ownership or any other mechanical effect which might arise due to the definition of the Amihud liquidity measure.

In the remainder of this section we now show that our results are also not dependent on the specification of the first stage liquidity covariance estimation procedure. We re-estimate β_{HI} in a variety of ways and report the results of second-stage tests of our main hypothesis [equation (2)] using the variety of first-stage β_{HI} estimates. These results are reported in columns 2 through 9 of Table 4.10. In the first approach, instead of using value-weighted portfolio liquidity to determine β_{HI}, we regress the individual stock liquidity measure on equal-weighted market and high mutual fund ownership portfolio liquidity after including the standard controls. Consistent with our results using value weighted portfolio liquidity, we find a very strong positive relation between the high mutual fund liquidity beta and mutual fund ownership. In this case, the coefficient is more than twice as large as the coefficient using value-weighted portfolio liquidity (2.063 in Table 4.10, column 2, compared to 0.836 in column 2 of Table 4.9). In the second approach, we employ our standard time series estimation procedure (model 1) but now follow Chordia, Roll, and Subrahmanyam (2000) and also use sum betas in the second stage, which equal β_{HI} plus the betas on the lead and lag values of the high mutual fund ownership (and similarly for the market beta). The results, reported in column 3 of Table 4.10, are consistent with our previous results. Next, the liquidity of stocks belonging to the same industry would be expected to comove more strongly with each other than with stocks not in the industry. Thus, in our third approach, we include industry-level measures in the first stage liquidity covariance estimation in two ways. The results reported in the fourth and fifth column of Table 4.10 use a β_{HI} estimated after controlling for the covariation between the firm's liquidity and that of a portfolio of stocks in its industry (identified by two-digit SIC code). In column 4 we use β_{HI} on the typical high $mfown$ portfolio, but we also control for liquidity covariation with stocks in the same industry by including lead, lag, and contemporaneous changes in the value-weighted industry portfolio liquidity. In column 5, we use a similar β_{HI} but additionally add the lead, lag, and contemporaneous return of the value weighted industry portfolio. In both cases, our measure of commonality in liquidity in high mutual fund ownership stocks, β_{HI}, has a positive and significant relationship with mfown. In columns 6 and 7 we use only one liquidity portfolio in the time series estimation. First, we remove the high mutual fund ownership portfolio (and its returns) and estimate a covariance with only the market portfolio. In column 7 we do the same using only a high mutual fund ownership portfolio. Not surprisingly, we find a positive relationship in the second stage between $mfown$ and β_{mkt}, and a positive but much stronger relationship between $mfown$ and β_{HI}. In column 8 we revert to the standard first stage portfolios and control variables used in the earlier tables. However, we now employ a different liquidity calculation to address the concern that changes in illiquidity might be over-differenced. Thus, as suggested by Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes (2009), we use a quasi-differencing method. Instead of using differences in logs of Amihud's illiquidity ratio we use the difference from a 5 day moving average. We find results that are similar to those from our main specification.

Finally we generate a portfolio of randomly selected stocks and include it instead of the portfolio of high mutual fund ownership stocks. Specifically, we randomly choose 25% of the stocks in each quarter and compute a value-weighted change in daily liquidity for this portfolio (and its returns). We then use liquidity betas on this portfolio as independent variable in our regression models. As expected, results in column 9 show that the liquidity beta on randomly selected stocks' liquidity in this placebo regression is not at all related to mutual fund ownership.

4.6 Conclusion

We hypothesize that correlated trading among investors in a stock is an important explanation for commonality in liquidity across stocks. Using data on mutual fund ownership and stock liquidity from NYSE and AMEX stocks for the period 1980 to 2008, we find evidence that suggests mutual funds are an important factor in explaining commonality in liquidity. We use a two-step process similar to the one suggested in Coughenour and Saad (2004) by first regressing a stock's liquidity on the liquidity of two portfolios: a market portfolio and a portfolio consisting of stocks with high mutual fund ownership. This regression results in two liquidity betas: a high mutual fund ownership portfolio liquidity beta and a market portfolio liquidity beta. In the second step, we examine the relation between the high mutual fund ownership liquidity beta and the extent to which a stock is owned by mutual funds. We find that mutual fund liquidity betas are about twice as large for stocks with high mutual fund ownership as for those with low mutual fund ownership. We also find that this result is not driven by time trends in commonality and mutual fund ownership or by stock characteristics such as firm size, liquidity levels, or other unobservable stock characteristics that might jointly determine systematic liquidity and mutual fund ownership.

We also expect the relation between commonality in liquidity and mutual fund ownership to be stronger in circumstances with greater mutual fund trading and our results support that hypothesis. We find that the commonality in liquidity is stronger in stocks that are owned by mutual funds with high turnover ratios. We also find that the commonality is greater during periods of negative or extreme aggregate mutual fund flows. Further, we find a strong positive relation between changes in aggregate mutual fund ownership and a stock's mutual fund liquidity beta.

Overall our results suggest that – in addition to the supply-side explanations for commonality in liquidity found in earlier studies (e.g., Coughenour and Saad, 2004; Comerton-Forde, Hendershott, Jones, Moulton and Seasholes, 2010) – demand-side factors, i.e., mutual fund ownership and particularly flow-induced trading, are important explanations as well. Thus, liquidity risk arises not only from the actions of market specialists, but also the investors in the stock. These results suggest that mutual fund trading may add to the risk of a stock, consistent with the findings of Sias (1996) that institutional investors contribute to a stock's volatility. Mutual fund managers might consider avoiding stocks with higher systematic liquidity risk, i.e., stocks whose ownership is dominated by other mutual funds, particularly if they are concerned about the effects of liquidity shocks hitting themselves in the form of investor flows. However, our results also suggest that this - at least in aggregate - is not possible, because mutual funds themselves give rise to much of the commonality in liquidity we observe.

In this paper we have selected mutual funds as a group of investors to examine for correlated trading and resulting commonality. Of course, this does not preclude the possibility that the correlated trading of investors such as hedge funds or other institutional investors might also give rise to commonality.

Table 4.1
Summary Statistics

This table reports summary statistics for select variables. Panel A reports statistics for the full sample of stock-quarters over the 1980-2008 period. *mfown* is the number of shares owned by mutual funds scaled by shares outstanding. *firm size* is the market value of the stock at the end of the quarter. *illiq(avg)* is the average over the quarter of the absolute value return scaled by dollar volume (in millions). *twmfown* is the total shares owned by mutual funds weighted by each fund's turnover, scaled by shares outstanding. Aggregate flows are the net dollar flows to or from all mutual funds in a quarter scaled by beginning of quarter total market value. Panel B reports means, standard deviations, and medians for subsamples of firms by *mfown* quartile ranked quarterly.

Panel A: Full Sample	N	Mean	Std Dev	Min	Max	Median
firm size (millions)	120,413	4270	16052	2	571197	897
illiq(avg)	120,413	0.08	0.3	< 0.001	215.74	0.008
mfown	120,413	0.13	0.1	0	0.88	0.10
twmfown	66,598	0.10	0.08	0	0.78	0.08
aggregate flows (% of mkt cap)	114	0.65%	0.73%	-3.05%	2.83%	0.65%

	mfown (ranked quarterly)			
Panel B: By mfown quartile	LO	2	3	HI
	Mean, (Std dev), Median			
firm size (millions)	3168	6686	4400	2821
	(14938)	(22869)	(11802)	(6487)
	401	1079	1199	1044
illiq(avg)	0.19	0.06	0.04	0.04
	(0.54)	(0.22)	(0.15)	(0.14)
	0.04	0.006	0.004	0.004
mfown	0.04	0.10	0.15	0.23
	(0.03)	(0.06)	(0.07)	(0.11)
	0.03	0.10	0.16	0.24
twmfown	0.03	0.08	0.12	0.19
	(0.03)	(0.04)	(0.06)	(0.10)
	0.02	0.07	0.11	0.17

Table 4.2
Time Series Estimates of Liquidity Betas

This table reports summary statistics on liquidity betas with respect to a high mutual fund ownership portfolio and a market portfolio of NYSE and AMEX stocks. Panel A reports these statistics for representative quarters in the sample. In each quarter and for each firm, the daily change in the firm's illiquidity (Amihud measure) is regressed on the daily changes in the illiquidity measure for a portfolio of high mutual fund ownership stocks and a market portfolio as well as control variables.

$$\Delta illiq_{i,t} = \alpha_i + \beta_{mfown} * \Delta illiq_{mfown,t} + \beta_{mkt} * \Delta illiq_{mkt,t} + controls$$

where $\Delta illiq_{i,t} = log\left[\frac{illiq_{i,t}}{illiq_{i,t-q}}\right] = log\left[\frac{\frac{|r_{i,t}|}{vold_{i,t}}}{\frac{|r_{i,t-1}|}{vold_{i,t-1}}}\right]$. In each time series regression the stock's individual measure is removed from the market portfolio and the high *mfown* portfolio (when applicable). The left columns summarize the coefficient estimates for the high *mfown* liquidity portfolio, and the right columns summarize the market liquidity portfolio. In each quarter we record the average beta, the percent positive and percent significant at the 5% level, and we compute a t-statistic on the sample of beta estimates in that quarter. Panel A reports averages for representative quarters and Panel B reports averages over 5 year periods and the full sample.

Panel A: Representative quarters

		HI *mfown* portfolio								Market portfolio						
	R^2	β_{HI}	%pos	%sig	tstat	size	mfown	illiq(avg)	#stocks	β_{mkt}	%pos	%sig	tstat	size	mfown	illiq(avg)
19802	0.30	0.22	54%	6%	3.39	543	0.09	0.083	301	0.20	56%	6%	3.86	878	0.04	0.126
19803	0.29	0.27	58%	6%	5.96	565	0.09	0.093	293	0.32	58%	8%	7.39	982	0.04	0.129
19804	0.30	0.57	67%	9%	11.97	637	0.09	0.084	275	0.15	54%	8%	3.19	1008	0.04	0.121
.																
19951	0.29	0.31	62%	6%	6.50	2753	0.22	0.010	333	0.23	54%	8%	3.92	3973	0.13	0.038
19952	0.29	0.10	51%	6%	1.89	2865	0.23	0.010	333	0.37	59%	8%	7.14	4108	0.14	0.036
19953	0.29	0.42	58%	9%	7.08	2945	0.22	0.012	340	0.31	59%	8%	5.68	4236	0.14	0.040
19954	0.31	0.51	64%	9%	10.20	3096	0.23	0.012	364	0.30	59%	7%	6.47	4353	0.14	0.042
.																
20081	0.32	0.29	60%	8%	8.02	3764	0.36	0.016	317	0.36	63%	12%	10.95	7869	0.22	0.091
20082	0.31	0.40	60%	10%	9.47	3060	0.36	0.010	309	0.38	62%	9%	9.55	7465	0.23	0.090
20083	0.31	0.19	55%	8%	5.20	1872	0.37	0.019	282	0.47	63%	13%	12.36	5807	0.24	0.138

Panel B: Five-year quarterly averages and full sample

1980-85	0.29	0.32	58%	7%	6.18	714	0.09	0.091	283	0.23	56%	7%	4.46	1201	0.05	0.120
1986-90	0.31	0.34	59%	8%	6.35	1670	0.11	0.046	254	0.22	57%	7%	4.46	2509	0.06	0.063
1991-95	0.30	0.30	58%	7%	5.44	2267	0.18	0.022	311	0.31	58%	8%	5.73	3487	0.11	0.045
1996-00	0.28	0.26	57%	6%	5.68	4953	0.26	0.018	410	0.24	56%	7%	5.55	5874	0.15	0.054
2001+	0.29	0.33	60%	8%	8.91	4023	0.33	0.012	341	0.33	61%	9%	9.14	6831	0.20	0.074
1980-08	0.29	0.31	58%	7%	6.66	2813	0.20	0.036	321	0.27	58%	8%	6.17	4204	0.12	0.073

Table 4.3
Liquidity Betas Sorted by Firm Characteristics

Panel A presents mutual fund and market liquidity betas sorted by firm characteristics. At the end of each quarter we sort stocks into quartiles based on *mfown*, *firm size*, or *illiq(avg)*. For each quartile we report the average β_{HI} and β_{mkt} measured over the subsequent quarter. Panel B presents dependent sorts. First we sort on *firm size* or *illiq(avg)* each quarter, then within each bin we sort on *mfown*. All t-statistics are on the difference in sample averages paired by quarter.

Panel A: One way sorts

	Average β_{HI}						Average β_{mkt}					
mfown	Lo	2	3	Hi	Hi - Lo	H-L tstat	Lo	2	3	Hi	Hi - Lo	H-L tstat
	0.20	0.28	0.35	0.40	0.20	(12.22)	0.24	0.33	0.29	0.24	0.00	(-0.49)
firm size	Lo	2	3	Hi	Hi - Lo	H-L tstat	Lo	2	3	Hi	Hi - Lo	H-L tstat
	0.23	0.33	0.38	0.29	0.06	(3.47)	0.09	0.20	0.27	0.53	0.44	(24.45)
illiq(avg)	Lo	2	3	Hi	Hi - Lo	H-L tstat	Lo	2	3	Hi	Hi - Lo	H-L tstat
	0.31	0.37	0.34	0.21	-0.09	(-5.90)	0.51	0.29	0.19	0.10	-0.41	(-22.97)

Panel B: Dependent sorts - First on size or illiq(avg) then on mfown

		Average β_{HI}: mfown							Average β_{mkt}: mfown					
		Lo	2	3	Hi	Hi - Lo	H-L tstat		Lo	2	3	Hi	Hi - Lo	H-L tstat
firm size	Small	0.18	0.28	0.26	0.27	0.09	(2.33)	Small	0.09	0.06	0.10	0.12	0.03	(1.02)
	2	0.22	0.27	0.36	0.42	0.20	(6.63)	2	0.23	0.22	0.20	0.18	-0.05	(-2.68)
	3	0.27	0.33	0.41	0.45	0.18	(6.46)	3	0.28	0.32	0.24	0.26	-0.02	(-0.98)
	Big	0.16	0.24	0.33	0.41	0.25	(9.48)	Big	0.61	0.61	0.50	0.40	-0.21	(-6.79)
		Lo	2	3	Hi	Hi - Lo	H-L tstat		Lo	2	3	Hi	Hi - Lo	H-L tstat
illiq(avg)	Lo	0.14	0.24	0.33	0.43	0.29	(12.07)	Lo	0.68	0.61	0.48	0.37	-0.31	(-9.78)
	2	0.27	0.31	0.40	0.44	0.17	(5.92)	2	0.35	0.35	0.25	0.25	-0.10	(-4.30)
	3	0.27	0.31	0.37	0.42	0.15	(4.91)	3	0.20	0.20	0.18	0.17	-0.03	(-2.67)
	Hi	0.17	0.26	0.24	0.24	0.07	(1.65)	Hi	0.11	0.06	0.12	0.13	0.02	(0.85)

Table 4.4
Relation Between Liquidity Commonality and Mutual Fund Ownership

This table reports results from the following Pooled OLS regression using alternate specifications:

$$\beta_{HI,i,t} = a + b_1 * mfown_{i,t-1} + b_2 * ln(firmsize_{i,t-1}) + b_3 * illiq(avg)_{i,t-1} + \varepsilon_{i,t}$$

where β_{HI} is estimated as in equation (1). *mfown* and *ln(firm size)* are measured at the end of the previous quarter. *illiq(avg)* is the firm's average daily Amihud (2002) illiquidity measure over the previous quarter. Panel A uses the standard measure of *mfown* and Panel B uses a dummy equal to 1 if *mfown* is in the top quartile in a given quarter, 0 otherwise. Quarter dummies are included in all regressions. Standard errors are clustered by stock.

Panel A	(1)	(2)	(3)	(4)	(5)
mfown	0.896***	0.838***	0.457***	0.557***	1.009***
	(14.73)	(13.12)	(4.58)	(5.33)	(9.23)
ln(firm size)		-0.0021	0.0187**	-0.0053	1.75e-05
		(-0.56)	(1.97)	(-1.10)	(0.00)
illiq(avg)		-0.0890***	-0.0529**	-0.1030***	-0.0954***
		(-4.75)	(-2.23)	(-5.50)	(-2.78)
Observations	120413	120413	120413	120413	120413
R^2	0.012	0.012	0.055	0.002	0.002

Panel B					
mfown (dummy)	0.127***	0.120***	0.0431***	0.120***	0.118***
	(11.37)	(10.69)	(3.09)	(9.06)	(9.45)
ln(firm size)		0.0037	0.0231**	0.0036	0.0030
		(0.97)	(2.44)	(0.73)	(0.69)
illiq(avg)		-0.106***	-0.0541**	-0.102***	-0.117***
		(-5.59)	(-2.27)	(-5.38)	(-3.37)
Observations	120413	120413	120413	120413	120413
R^2	0.011	0.011	0.055	0.002	0.002
Time effects	Y	Y	Y		
Firm effects			Y		
Time clusters				Y	
Firm clusters	Y	Y	Y	Y	
Fama MacBeth					Y

Table 4.5
Relation Between Liquidity Commonality and Mutual Fund Ownership: Subsample Analysis

This table reports results from the following Pooled OLS regression using various sub-samples based on size, average illiquidity, and time:

$$\beta_{HI,i,t} = a + b_1 * mfown_{i,t-1} + b_2 * ln(firmsize_{i,t-1}) + b_3 * illiq(avg)_{i,t-1} + \varepsilon_{i,t}$$

where β_{HI} is estimated as in equation (1). *mfown* and *ln(firm size)* are measured at the end of the previous quarter. *illiq(avg)* is the firm's average daily Amihud (2002) illiquidity measure over the previous quarter. Panels A and C report results of regressions for size and illiquidity quartiles. Panels B and D report results of regressions for five year subperiods and for up and down markets separately, where up(down) market periods are quarters in which the market return was positive(negative). Panels A and B use the standard measure of *mfown*, and Panels C and D use a dummy equal to 1 if *mfown* is in the top quartile in a given quarter, 0 otherwise. Quarter dummies are included in all regressions. Standard errors are clustered by stock.

	size				illiq(avg)			
Panel A	Lo	2	3	Hi	Lo	2	3	Hi
mfown	0.155	0.738***	0.761***	1.008***	1.016***	0.668***	0.659***	0.151
	(1.11)	(6.81)	(6.41)	(6.90)	(7.12)	(5.31)	(5.96)	(1.04)
ln(firm size)	0.0513***	0.0301	-0.0336	-0.0733***	-0.0800***	-0.0344*	0.0108	0.0192
	(3.18)	(0.99)	(-1.20)	(-5.40)	(-6.44)	(-1.90)	(0.70)	(1.54)
illiq(avg)	-0.0334	-0.304	0.347	-1.032	-20.46***	-4.011*	1.038	-0.0402*
	(-1.57)	(-1.43)	(1.21)	(-0.88)	(-2.89)	(-1.76)	(1.41)	(-1.93)
Observations	30057	30120	30150	30086	30057	30120	30150	30086
R^2	0.010	0.023	0.018	0.018	0.018	0.021	0.021	0.010

Panel B	1980-85	86-90	91-95	96-00	2001+	Down mkt	Up mkt
mfown	1.095***	1.487***	1.187***	0.349***	1.006***	0.950***	0.785***
	(3.49)	(5.00)	(7.61)	(2.85)	(12.64)	(9.63)	(10.43)
ln(firm size)	0.0161*	-0.0007	0.0038	0.0004	-0.0049	0.0095*	-0.0073*
	(1.65)	(-0.07)	(0.62)	(0.07)	(-0.84)	(1.69)	(-1.70)
illiq(avg)	-0.0763	-0.0662	-0.0991***	-0.0465	-0.0889***	-0.0674***	-0.101***
	(-1.58)	(-1.01)	(-2.73)	(-0.81)	(-3.86)	(-2.66)	(-3.77)
Observations	21915	15885	51717	26587	38348	37325	83088
R^2	0.007	0.010	0.009	0.011	0.018	0.016	0.011

	size				illiq(avg)			
Panel C	Lo	2	3	Hi	Lo	2	3	Hi
mfown (dummy)	0.0108	0.109***	0.0971***	0.138***	0.131***	0.0889***	0.0822***	0.0154
	(0.37)	(5.40)	(5.05)	(6.20)	(6.32)	(4.64)	(3.92)	(0.48)
ln(firm size)	0.0551***	0.0348	-0.0311	-0.0781***	-0.0881***	-0.0408**	0.0039	0.0199
	(3.49)	(1.14)	(-1.10)	(-5.72)	(-7.06)	(-2.28)	(0.25)	(1.59)
illiq(avg)	-0.0345	-0.372*	0.184	-1.462	-22.71***	-4.495**	0.835	-0.0425**
	(-1.63)	(-1.75)	(0.65)	(-1.21)	(-3.21)	(-1.97)	(1.13)	(-2.06)
Observations	30057	30120	30150	30086	30057	30120	30150	30086
R^2	0.010	0.023	0.017	0.018	0.017	0.020	0.020	0.009

Panel D	1980-85	86-90	91-95	96-00	2001+	Down mkt	Up mkt
mfown (dummy)	0.0728***	0.116***	0.116***	0.0770***	0.166***	0.140***	0.111***
	(2.79)	(4.09)	(6.72)	(3.25)	(9.39)	(7.56)	(8.38)
ln(firm size)	0.0159	-0.0020	0.0048	0.0020	0.0067	0.0163***	-0.0020
	(1.62)	(-0.20)	(0.77)	(0.30)	(1.17)	(2.90)	(-0.45)
illiq(avg)	-0.0844*	-0.0808	-0.110***	-0.0551	-0.113***	-0.0852***	-0.116***
	(-1.75)	(-1.23)	(-3.02)	(-0.95)	(-4.80)	(-3.37)	(-4.32)
Observations	21915	15885	51717	26587	38348	37325	83088
R^2	0.007	0.009	0.008	0.011	0.015	0.015	0.010

Table 4.6
Relation Between Liquidity Commonality and Turnover-weighted Mutual Fund Ownership

This table reports results from a pooled OLS regression using a turnover weighted measure of mutual fund ownership. Specifically we compute for each stock i at quarter t,

$$twmfown_{i,t} = \frac{\sum_{j=1}^{J} sharesowned_{j,i,t} * turnover_{j,t}}{shrout_{i,t}}$$

where $sharesowned_{j,i,t}$ is the ownership of fund j in stock i at end of quarter t from CDA/Spectrum and $turnover_{j,t}$ is the turnover reported by CRSP for fund j over quarter t. Results are reported for the following regression using the subsample in which the turnover variable is available quarterly from CRSP (1999+):

$$\beta_{HI,i,t} = a + b_1 * twmfown_{i,t-1} + b_2 * mfown_{i,t-1} + b_3 * ln(firmsize_{i,t-1}) + b_4 * illiq(avg)_{i,t-1} + \varepsilon_{i,t}.$$

Model (1) includes *twmfown* and for comparison model (2) includes the standard (unweighted) *mfown* over the same sample for which turnover is available (1999+), and model (3) includes both variables. To facilitate comparison of coefficients, the last three models repeat the first three but use standardized values of *twmfown* and *mfown*. Quarter dummies are included but not reported. Standard errors are clustered by stock.

				standardized variables		
	(1)	(2)	(3)	(4)	(5)	(6)
twmfown	1.331***		1.152***	0.112***		0.0972***
	(15.45)		(8.31)	(15.45)		(8.31)
mfown		0.925***	0.185		0.0935***	0.0188
		(12.65)	(1.60)		(12.65)	(1.60)
ln(firm size)	-0.0026	-0.0031	-0.0035	-0.0026	-0.0031	-0.0035
	(-0.54)	(-0.60)	(-0.72)	(-0.54)	(-0.60)	(-0.72)
illiq(avg)	-0.0750***	-0.0787***	-0.0733***	-0.0750***	-0.0787***	-0.0733***
	(-3.39)	(-3.55)	(-3.31)	(-3.39)	(-3.55)	(-3.31)
Observations	48907	48907	48907	48907	48907	48907
R^2	0.021	0.020	0.021	0.021	0.020	0.021

Table 4.7
Relation Between Liquidity Commonality and Mutual Fund Ownership Conditional on Flows

This table reports results from a Pooled OLS regression of β_{HI} on *mfown* conditional on fund flows. We define dummy variables based on one of three measures of flows; aggregate net flows, aggregate inflows, or aggregate outflows in each quarter. All aggregate flows are scaled by total US market capitalization. Flows are measured contemporaneously with β_{HI}. The dummy variable *hiabsflow* equals one if aggregate net flows are in either the highest 10% or lowest 10%, zero otherwise. *negnetflow* equals one if aggregate net flows are negative (outflows) for that quarter and zero otherwise. We then define two dummy variables using inflows and outflows aggregated separately in each quarter. *hiinflow* equals one for the top 25% of quarters of inflows scaled by market cap, and *hioutflow* equals one for the top 25% of quarters of outflows scaled by market cap. Models (1)-(4) use the full sample. Models (5)-(8) show the effect of *mfown* within subsamples defined by aggregate net flows. Quarter dummies are included but not reported. Standard errors are clustered by stock.
In Panel B we first run 115 cross sectional regressions of β_{HI} on *mfown* and control for size and liquidity. Then we regress the time series of *mfown* coefficients on aggregate flows and the square of aggregate flows in order to test for a U-shaped relationship.

	Full sample				Subsamples:Agg flows as % of US mkt cap			
Panel A					$< 0\%$	0 to 0.5%	0.5 to 1%	$> 1\%$
	(1)	(2)	(3)	(4)	(5)	(6)		
mfown	0.765***	0.762***	1.174***	0.852***	0.710***	0.935***		
	(11.13)	(11.33)	(7.97)	(7.04)	(8.01)	(7.14)		
hiabsflow * mfown	0.395***							
	(3.12)							
negnetflow * mfown		0.575***						
		(3.91)						
ln(firm size)	-0.0019	-0.0018	-0.0005	-0.0083	-0.0023	0.0037		
	(-0.50)	(-0.47)	(-0.062)	(-1.23)	(-0.47)	(0.52)		
illiq(avg)	-0.0880***	-0.0880***	-0.106**	-0.135***	-0.0960***	-0.0157		
	(-4.70)	(-4.70)	(-2.14)	(-3.62)	(-3.53)	(-0.54)		
Observations	120413	120413	16873	23900	53604	26036		
R^2	0.012	0.012	0.012	0.012	0.013	0.008		

Panel B

Dependent variable: Coefficient on *mfown*	
aggflows	-1.04**
	(-2.09)
aggflows^2	0.57***
	(3.07)
Constant	1.28***
	(4.95)
Observations	115
R-squared	0.11

Table 4.8
Relation Between Liquidity Commonality and Changes in Mutual Fund Ownership

This table reports results of a Pooled OLS regression of β_{HI} at time t on the absolute value of the change in *mfown* from $t-1$ to t, lagged firm size and lagged average illiquidity:

$$\beta_{HI,i,t} = a + b_1 * |\Delta_{t-1,t} mfown_i| + b_2 * ln(firmsize_{i,t-1}) + b_3 * illiq(avg)_{i,t-1} + \varepsilon_{i,t}.$$

In model (2) we replace the absolute change in mutual fund ownership with a dummy variable set to one if the absolute change is in the top quartile in that quarter. Quarter dummies are included but not reported. Standard errors are clustered by stock.

	(1)	(2)
$\lvert\Delta_{t-1,t} mfown\rvert$	1.029***	
	(4.620)	
$\lvert\Delta_{t-1,t} mfown\rvert$ (dummy)		0.0399***
		(9.265)
ln(firm size)	0.0002	0.0016
	(0.047)	(0.378)
illiq(avg)	-0.137***	-0.116***
	(-4.412)	(-3.779)
Observations	105312	105312
R^2	0.011	0.011

Table 4.9
Robustness Tests: Controlling for Return and Volatility Covariation

This table reports results from Pooled OLS regressions of β_{HI} on *mfown* with additional controls. Panel A reports results controlling for commonality in returns and Panel B reports results controlling for commonality in volatility. The first model repeats the standard regression of β_{HI} on mutual fund ownership (as in models (1) and (2) of Table 4.4) and includes as a control variable the beta estimate between the firm return and the value-weighted return on the high mutual fund ownership portfolio estimated contemporaneously with the liquidity beta. Models (2)-(5) run the standard regression on cross-sectional subsamples sorted by the return beta. Model (6) runs the standard regression, but controls for return covariation in the first stage. Specifically, the dependent variable is a liquidity beta estimated in a time series regression that controls for firm returns and the return on the high mutual fund ownership portfolio. We repeat this analysis in Panel B, substituting returns-squared for returns, as a proxy for volatility.

Panel A: Controlling for covariation in returns

		mutual fund return beta subsamples				1st stage control
	full	Lo	2	3	Hi	for returns
	(1)	(2)	(3)	(4)	(5)	(6)
mfown	0.706***	0.619***	0.716***	0.516***	0.620***	0.806***
	(11.25)	(5.34)	(5.89)	(4.45)	(5.44)	(12.08)
ln(firm size)	0.0009	-0.0260***	-0.0126*	0.0174**	0.0468***	0.00125
	(0.25)	(-4.67)	(-1.91)	(2.54)	(6.73)	(0.32)
illiq(avg)	-0.0807***	-0.0641***	-0.121***	-0.0950	-0.0709**	-0.0707***
	(-4.33)	(-2.64)	(-3.06)	(-1.63)	(-2.09)	(-3.33)
mutual fund return beta	0.051***					
	(17.42)					
Observations	120413	30057	30120	30150	30086	120413
R^2	0.015	0.016	0.015	0.016	0.021	0.011

Panel B: Controlling for covariation in returns-squared

		mutual fund return^2 beta subsamples				1st stage control
	full	Lo	2	3	Hi	for ret squared
	(1)	(2)	(3)	(4)	(5)	(6)
mfown	0.830***	0.673***	0.839***	0.638***	0.671***	0.800***
	(13.01)	(6.16)	(7.07)	(5.32)	(5.64)	(11.93)
ln(firm size)	-0.0020	-0.0145**	-0.0230***	0.0117*	0.0352***	0.00174
	(-0.52)	(-2.32)	(-3.48)	(1.87)	(5.22)	(0.44)
illiq(avg)	-0.0876***	-0.0663***	-0.139**	-0.157***	-0.0627*	-0.0948***
	(-4.69)	(-2.72)	(-2.27)	(-3.95)	(-1.88)	(-4.56)
mutual fund	0.0022***					
phantomasreturn^2 beta	(4.84)					
Observations	120413	30057	30120	30150	30086	120413
R^2	0.012	0.014	0.015	0.015	0.022	0.012

Table 4.10
Robustness Tests: Alternate Measures of Liquidity Betas

This table reports the results of Pooled OLS regressions of β_{HI} on mutual fund ownership (as in the first two models of Table 4.4) using alternate measures of liquidity betas. In model (1) the dependent variable is the liquidity beta estimate on an equal-weighted portfolio of high *mfown* stocks instead of a value-weighted portfolio. In model (2) the dependent variable is a sum beta that equals β_{HI} plus the betas on lead and lag values of the high *mfown* portfolio (measured in the standard way). In model (3) we use β_{HI} on the typical high *mfown* portfolio, but we also control for liquidity covariation with stocks in the same industry (lead, lag, and contemporaneous changes in the industry portfolio as identified by two-digit SIC code). Model (4) uses β_{HI} from a similar time series regression as in model (3), but we also include contemporaneous, lead and lag returns on the high *mfown* portfolio as well as those on the industry portfolio. Models (5) and (6) use only one portfolio in the time series beta estimation, the market portfolio and the high *mfown* portfolio respectively. Model (7) reports results using changes in liquidity from a five day moving average (as opposed to a first difference). Specifically we compute the change in illiquidity as the log of the ratio of Amihud's illiquidity measure at day t to the average of this measure of the previous five trading days. Model (8) uses turnover instead of Amihud's illiquidity measure. The last model uses the beta on a portfolio of randomly selected stocks. Specifically we choose 25% of the stocks in each quarter and compute a value-weighted change in daily liquidity for this random portfolio. Quarter dummies are included in all regressions and standard errors are clustered by stock.

	(1) turnover	(2) equal weight	(3) sum betas	(4) industry controls	(5) ind and ret controls	(6) β_{mkt} only	(7) β_{HI} only	(8) quasi-differencing	(9) random MO port
mfown	1.830***	2.063***	0.810***	0.761***	0.760***	0.314***	0.504***	0.721***	0.0611
	(18.96)	(15.19)	(6.87)	(11.67)	(9.49)	(7.79)	(12.41)	(12.58)	(1.46)
ln(firm size)	-0.0712***	0.0491***	-0.0176***	-0.0057	-0.0003	0.114***	0.101***	-0.0046	0.0027
	(-12.10)	(6.51)	(-2.73)	(-1.52)	(-0.07)	(43.78)	(40.11)	(-1.45)	(1.20)
illiq(avg)	-0.128***	-0.0903**	-0.0694*	-0.0790***	0.0090	-0.0039	-0.0277***	-0.0778***	0.0079
	(-2.71)	(-2.35)	(-1.95)	(-3.92)	(0.12)	(-0.39)	(-2.58)	(-3.93)	(0.52)
Observations	120413	120413	120413	120114	120114	120413	120413	120413	120413
R^2	0.017	0.007	0.006	0.008	0.005	0.075	0.066	0.014	0.012

Bibliography

[1] Viral V. Acharya and Lasse Heje Pedersen. Asset pricing with liquidity risk. *Journal of Financial Economics*, 77(2):375–410, 2005.

[2] Yakov Amihud. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1):31–56, 2002.

[3] Yakov Amihud and Haim Mendelson. Liquidity, maturity, and the yields on u.s. treasury securities. *The Journal of Finance*, 46(4):1411–1425, 1991.

[4] Andrew Ang, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299, 2006.

[5] M. Anton and C. Polk. Connected stocks. *Working Paper*, 2010.

[6] Christopher Avery and Peter Zemsky. Multidimensional uncertainty and herd behavior in financial markets. *American Economic Review*, 88:724–748, 1998.

[7] Abhijit V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, 107(3):797–817, 1992.

[8] Nicholas Barberis and Andrei Shleifer. Style investing. *Journal of Financial Economics*, 68(2):161–199, 2003.

[9] Nicholas Barberis, Andrei Shleifer, and Jeffrey Wurgler. Comovement. *Journal of Financial Economics*, 75(2):283–317, 2005.

[10] Geert Bekaert and Guojun Wu. Asymmetric volatility and risk in equity markets. *The Review of Financial Studies*, 13(1):1–42, 2000.

[11] Azi Ben-Rephael, Shmuel Kandel, and Avi Wohl. The price pressure of aggregate mutual fund flows. *Journal of Financial and Quantitative Analysis*, 46(02):585–603, 2011.

[12] James A. Bennett, Richard W. Sias, and Laura T. Starks. Greener pastures and the impact of dynamic institutional preferences. 16(4):1203–1238, 2003.

[13] Jonathan B. Berk and Richard C. Green. Mutual fund flows and performance in rational markets. *The Quarterly Journal of Economics*, 112(6):1269–1295, 2004.

[14] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *The Journal of Political Economy*, 100(5):992–1026, 1992.

[15] Michael J. Brennan and Avanidhar Subrahmanyam. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics*, 41(3):441–464, 1996.

[16] Paul Brockman and Dennis Y. Chung. Commonality in liquidity: Evidence from an order-driven market structure. *Journal of Financial Research*, 25(4):521–539, 2002.

[17] Mark M. Carhart. On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82, 1997.

[18] Judith Chevalier and Glenn Ellison. Career concerns of mutual fund managers. *The Quarterly Journal of Economics*, 114(2):389–432, 1999.

[19] Tarun Chordia, Richard Roll, and Avanidhar Subrahmanyam. Commonality in liquidity. *Journal of Financial Economics*, 56(1):3–28, 2000.

[20] Tarun Chordia, Richard Roll, and Avanidhar Subrahmanyam. Order imbalance, liquidity, and market returns. *Journal of Financial Economics*, 65(1):111–130, 2002.

[21] Tarun Chordia, Richard Roll, and Avanidhar Subrahmanyam. Why has trading volume increased? *Working Paper*, 2009.

[22] Tarun Chordia, Asani Sarkar, and Avanidhar Subrahmanyam. An empirical analysis of stock and bond market liquidity. *Review of Financial Studies*, 18(1):85–129, 2005.

[23] Randy Cohen, C. Polk, and B. Silli. Best ideas. *Working Paper*, 2009.

[24] Carole Comerton-Forde, Terrence Hendershott, Charles M. Jones, Pamela C. Moutlon, and Mark S. Seaholes. Time variation in liquidity: The role of market-maker inventories and revenues. *The Journal of Finance*, 65(1):295–331, 2010.

[25] Jay F. Coughenour and Mohsen M. Saad. Common market makers and commonality in liquidity. *Journal of Financial Economics*, 73(1):37–69, 2004.

[26] Joshua Coval and Erik Stafford. Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, 86(2):479–512, 2007.

[27] K. J. Martijn Cremers and Antti Petajisto. How active is your fund manager? a new measure that predicts performance. *The Review of Financial Studies*, 22(9):3329–3365, 2009.

[28] K. J. Martijn Cremers, Antti Petajisto, and Eric Zitzewitz. Should benchmark indicies have alpha? revisiting performance evaluation. *Working Paper*, 2010.

[29] Zhi Da, Pengjie Gao, and Ravi Jagannathan. Informed trading, liquidity provision, and stock selection by mutual funds. (14609), 2008.

[30] Kent Daniel, Mark Grinblatt, Sheridan Titman, and Russ Wermers. Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance*, 52(3):1035–1058, 1997.

[31] Amil Dasgupta and Andrea Prat. Information aggregation in financial markets with career concerns. *Journal of Economic Theory*, 143(1):83 – 113, 2008.

[32] Amil Dasgupta, Andrea Prat, and Michela Verardo. The price impact of institutional herding. *Review of Financial Studies*, 24(3):892–925, 2011.

[33] Nishant Dass, Massimo Massa, and Rajdeep Patgiri. Mutual funds and bubbles: The surprising role of contractual incentives. *Review of Financial Studies*, 21(1):51–99, 2008.

[34] Andrea Devenow and Ivo Welch. Rational herding in financial economics. *European Economic Review*, 40(3-5):603 – 615, 1996.

[35] E. Eckbo and O. Norli. Pervasive liquidity risk. *Working Paper*, 2002.

[36] Eric G Falkenstein. Preferences for stock characteristics as revealed by mutual fund portfolio holdings. *Journal of Finance*, 51(1):111–35, 1996.

[37] Eugene F. Fama and Kenneth R. French. Industry costs of equity. *Journal of Financial Economics*, 43(2):153–193, 1997.

[38] Eugene F Fama and James D MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–36, 1973.

[39] Andrea Frazzini and Owen A. Lamont. Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics*, 88(2):299–322, 2008.

[40] Kenneth A. Froot, David S. Scharfstein, and Jeremy C. Stein. Herd on the street: Informational inefficiencies in a market with short-term speculation. *The Journal of Finance*, 47(4):1461–1484, 1992.

[41] Paul A. Gompers and Andrew Metrick. Institutional investors and equity prices. *The Quarterly Journal of Economics*, 116(1):229–259, 2001.

[42] Ruslan Y. Goyenko, Craig W. Holden, and Charles A. Trzcinka. Do liquidity measures measure liquidity? *Journal of Financial Economics*, 92(2):153–181, 2009.

[43] T. Clifton Green and Byoung-Hyoun Hwang. Price-based return comovement. *Journal of Financial Economics*, 93(1):37–50, 2009.

[44] Robin Greenwood. Excess comovement of stock returns: Evidence from cross-sectional variation in nikkei 225 weights. *Review of Financial Studies*, 21(3):1153–1186, 2008.

[45] Robin Greenwood and David Thesmar. Stock price fragility. *Working Paper*, 2009.

[46] Mark Grinblatt, Sheridan Titman, and Russ Wermers. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *The American Economic Review*, 85(5):pp. 1088–1105, 1995.

[47] Diane Del Guercio. The distorting effect of the prudent-man laws on institutional equity investments. *Journal of Financial Economics*, 40(1):31–62, 1996.

[48] Allaudeen Hameed, Wenjin Kang, and S. Viswanathan. Stock market declines and liquidity. *Journal of Finance*, 65(1):257–293, 2010.

[49] Joel Hasbrouck. Trading costs and returns for u.s. equities: Estimating effective costs from daily data. *Journal of Finance*, 64(3):1445–1477, 2009.

[50] Joel Hasbrouck and Duane J. Seppi. Common factors in prices, order flows, and liquidity. *Journal of Financial Economics*, 59(3):383–411, 2001.

[51] Terrence Hendershott and Mark Seasholes. Market predictablility and non-informational trading. *Working Paper*, 2009.

[52] David Hirshleifer, Avanidhar Subrahmanyam, and Sheridan Titman. Security analysis and trading patterns when some investors receive information before others. *The Journal of Finance*, 49(5):1665–1698, 1994.

[53] Gur Huberman and Dominika Halka. Systematic liquidity. *Journal of Financial Research*, 24(2):161–78, 2001.

[54] C.M. Jones. A century of stock market liquidity and trading costs. *Working Paper*, 2002.

[55] Marcin Kacperczyk, Clemens Sialm, and Lu Zheng. On the industry concentration of actively managed equity mutual funds. *The Journal of Finance*, 60(4):1983–2011, 2005.

[56] Marcin Kacperczyk, Clemens Sialm, and Lu Zheng. Unobserved actions of mutual funds. *The Review of Financial Studies*, 21(6):2379–2416, 2008.

[57] Avraham Kamara, Xiaoxia Lou, and Ronnie Sadka. The divergence of liquidity commonality in the cross-section of stocks. *Journal of Financial Economics*, 89(3):444–466, 2008.

[58] J. M. Keynes. The general theory of employment. *The Quarterly Journal of Economics*, 51(2):209–223, 1937.

[59] M. Khan, L. Kogan, and G. Serafeim. Mutual fund trading pressure: Firm-level stock price impact and timing of seos. *Working Paper*, 2009.

[60] Ajay Khorana. Top management turnover an empirical investigation of mutual fund managers. *Journal of Financial Economics*, 40(3):403–427, 1996.

[61] Robert A. Korajczyk and Ronnie Sadka. Pricing the commonality across alternative measures of liquidity. *Journal of Financial Economics*, 87(1):45–72, 2008.

[62] Alan Kraus and Hans R. Stoll. Parallel trading by institutional investors. *The Journal of Financial and Quantitative Analysis*, 7(5):2107–2138, 1972.

[63] Alok Kumar and Charles M.C. Lee. Retail investor sentiment and return comovements. *Journal of Finance*, 61(5):2451–2486, 2006.

[64] Josef Lakonishok, Andrei Shleifer, and Robert W. Vishny. The impact of institutional trading on stock prices. *Journal of Financial Economics*, 32(1):23–43, 1992.

[65] Josef Lakonishok, Andrei Shleifer, and Robert W. Vishny. Contrarian investment, extrapolation, and risk. *The Journal of Finance*, 49(5):1541–1578, 1994.

[66] M.T. Leary and M.R. Roberts. Do peer firms affect corporate financial policy? *Working Paper*, 2010.

[67] Kuan-Hui Lee. The world price of liquidity risk. *Journal of Financial Economics*, 99(1):136–161, 2011.

[68] J. Bradford de Long, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, 45(2):379–395, 1990.

[69] Massimo Massa. Mutual fund competition and stock market liquidity. *CEPR Working Paper*, 2004.

[70] Massimo Massa and Rajdeep Patgiri. Compensation and managerial herding: Evidence from the mutual fund industry. *Working Paper*, 2010.

[71] Massimo Massa and Ludovic Phalippou. Mutual fund competition and stock market liquidity. *CEPR Working Paper*, 2004.

[72] Whitney K. Newey and Kenneth D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708, 1987.

[73] Lubos Pastor and Robert F. Stambaugh. Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3):642–685, 2003.

[74] Mitchell A. Petersen. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies*, 22(1):435–480, 2009.

[75] Christo Pirinsky and Qinghai Wang. Does corporate headquarters location matter for stock returns? *The Journal of Finance*, 61(4):1991–2015, 2006.

[76] Joshua M. Pollet and Mungo Wilson. How does size affect mutual fund behavior? *The Journal of Finance*, 63(6):2941–2969, 2008.

[77] Lucasz Pomorski. Acting on the most valuable information: "best ideas" trades of mutual fund managers. *Working Paper*, 2009.

[78] Lucasz Pomorski. Follow the leader: Peer effects in mutual fund portfolio decisions. *Working Paper*, 2009.

[79] Andrea Prat and Amil Dasgupta. Financial equilibrium with career concerns. *Theoretical Economics*, 1(1):67–93, 2006.

[80] Canice Prendergast and Lars Stole. Impetuous youngsters and jaded old-timers: Acquiring a reputation for learning. *The Journal of Political Economy*, 104(6):1105–1134, 1996.

[81] David S. Scharfstein and Jeremy C. Stein. Herd behavior and investment. *The American Economic Review*, 80(3):465–479, 1990.

[82] Robert J. Shiller and John Pound. Survey evidence on diffusion of interest and information among investors. *Journal of Economic Behavior & Organization*, 12(1):47–66, 1989.

[83] Richard W. Sias. Volatility and the institutional investor. *Financial Analysts Journal*, 52(2):13–20, 1996.

[84] Richard W. Sias. Institutional herding. *The Review of Financial Studies*, 17(1):165–206, 2004.

[85] Richard W. Sias and Laura T. Starks. Return autocorrelation and institutional investors. *Journal of Financial Economics*, 46(1):103–131, 1997.

[86] Richard W. Sias and Laura T. Starks. Changes in institutional ownership and stock returns: Assessment and methodology. *Journal of Business*, 79(6):2869–2910, 2006.

[87] John Toner and Yuhai Tu. Flocks, herds, and schools: A quantitative theory of flocking. *Phys. Rev. E*, 58(4):4828–4858, 1998.

[88] Dimitri Vayanos. Flight to quality, flight to liquidity, and the pricing of risk. *NBER Working Papers*, 2004.

[89] Russ Wermers. Mutual fund herding and the impact on stock prices. *The Journal of Finance*, 54(2):581–622, 1999.

[90] Jeffrey Zwiebel. Corporate conservatism and relative compensation. *The Journal of Political Economy*, 103(1):1–25, 1995.

Vita

Andrew Wallace Koch was born in Waco, Texas on 17 September 1977, the son of Dr. Timothy W. Koch and Jan K. Koch. He received the Bachelor of Science degree in Engineering from Georgia Institute of Technology in 2001. He received a Masters in Business Administration from the University of Arizona in 2005.

Permanent address: 126 Blackburn Rd West
Irmo, SC 29063

This dissertation was typeset with LaTeX[†] by the author.

[†] LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.